A Model of Attention and Anticipation*

Joshua Tasoff[†] and Kristóf Madarász

UC Berkeley and LSE

JOB MARKET PAPER

November 12, 2009

Abstract

We develop a model in which people experience standard consumption utility, as well as anticipatory utility defined as the weighted sum of independently anticipated consumption "episodes" or "dimensions". The weights on these dimensions correspond to the attention that the person pays to the dimension. We assume attention on a dimension increases when expected consumption utility in the dimension differs from expected consumption utility under the default action or the prior belief. We show that the decision maker will pay more for information about dimensions with high expected consumption utility, and the willingness to pay may be negative when expected consumption utility is low. Additionally, when expected consumption utility is sufficiently low, but not when it is high, the decision maker will follow the default action even if it is suboptimal from a consumption standpoint. Furthermore, given the decision maker's current beliefs and preferences in a dimension, he will consume more in that dimension if he just received information. We then consider an advertisement application in which a monopolist decides whether to certifiably reveal the quality of various exogenous attributes of a good to a consumer who may choose to buy or not. There exists a sequential equilibrium for which the monopolist will not disclose information for attributes in which the consumer's utility with the highest quality good is sufficiently worse than not buying the good. Competition increases disclosure.

*We are deeply indebted to Botond Kőszegi for continual advice and guidance throughout all stages of this project and to Matthew Rabin for countless comments and suggestions. We would also like to thank Rosario Macera, Karen McKeown, Matthew R. Levy, Paul Heidhues, Stefano DellaVigna participants at the Berkeley Psychology and Economics Seminar and the Berkeley Psychology and Economics Non-lunch for many useful comments. The authors alone are responsible for any errors.

[†]Email: tasoff@econ.berkeley.edu (Corresponding Author).

1 Introduction

The growing literature on anticipatory utility assumes people experience pleasure and pain in the present from future consumption. Imagining one's future deteriorating health may cause immediate feelings of anxiety, while imagining one's next vacation may arouse immediate feelings of excitement. Anticipatory utility can affect individual decision making in a number of economically important ways: people may delay pleasurable experiences and expedite distasteful experiences, they may behave in a time inconsistent manner (Loewenstein, 1987), and they may have intrinsic information preferences (Caplin and Leahy, 2001).

A critical open issue in this line of research is to determine which future experiences a person anticipates in the present. Although most people can expect many happy and sad episodes in their future, at any given moment a person derives anticipatory utility from only a subset of future experiences. A person need not anticipate both their future deteriorating health and their next vacation simultaneously.

In this paper we explore the economic consequences of a person's ability to influence, through actions and information acquisition, the set of future anticipated episodes from which they derive anticipatory utility. The ability to influence this set creates a new class of unexplored incentives.

We posit that a person experiences more intense anticipatory utility from an expected future episode when the person directs more of his or her attention to that episode. Specifically, we assume that anticipatory utility is additively separable in I dimensions, and we define attention on a dimension as the weight placed on anticipatory utility in that dimension. Attention increases on a dimension in two situations. First, attention increases on a dimension when a person receives information that changes expected consumption utility in that dimension. Information about one's illness will make one think more of one's future health, and receiving information about Hawaii will make one think more about one's upcoming vacation. Second, attention increases on a dimension when the chosen action differs from the habitual action and produces different expected consumption utility in that dimension. When taking non-habitual actions to combat one's illness, one thinks about one's future health, and when packing one's bags for a leisurely vacation to Hawaii, one imagines being in Hawaii.

In Section 2 we review the economic literature on anticipatory utility and other related models.

In Section 3 we present a formal model of attention and anticipation. While our model requires substantial modeling judgment to apply it to specific economic environments, it does capture several important issues that are economically relevant. The model has three periods. The decision maker begins by choosing a "default". The default is not a fully contingent strategy but an action. This default is considered to be the action that requires no additional attention to enact. We interpret the default action as a habit, or as a pre-implemented effect that would occur in the future if one were to omit action at a later date. For example, a person may have the habit of eating fatty food and not taking heart medication. This would be the person's default. An investor's current portfolio would be his default since he can passively keep the status quo. Choosing any other portfolio would require attention. After establishing a default, the decision maker may choose information. For instance, he may choose to learn about his heart condition. He may then continue with his default or may choose a different action. Anticipatory utility is experienced and then consumption utility is experienced.

Consumption utility is standard expected utility defined over consequences z, and occurs at the end of period three. The separability assumption assigns "dimensions" to qualitatively different forms of consumption, or to different points in time. For example, leisure and health, or vacation next month versus vacation next year, may be represented by different dimensions. Psychologically, we interpret these dimensions as consumption experiences that one can think about independently of other consumption experiences. For example, most people can think of their next vacation a month away without thinking of their health ten years into the future. Arguably, even contemporaneous consumption can be attended to independently. A person may think about what clothes to wear tomorrow and anticipate the experience without thinking about what will be eaten tomorrow for lunch. Total consumption utility is the sum over all the I dimensions indexed by i, $\sum_{i=1}^{I} u_i(z)$.

Anticipatory utility is weighted expected consumption utility, and it occurs at the end of period two. Attention increases on a dimension only indirectly when information changes expected consumption utility in dimension i, and actions change expected consumption utility from the default expected consumption utility in dimension i. Total attention $A > 0$ is limited. More attention on future consumption is less attention on something else, such as less attention on the present. We assume there is a constant forgone utility of attention u_p. This is the hedonic experience that would have been felt if attention were not on the future consumption dimension. To illustrate, this may be the utility from paying attention to the Johnny Depp movie on television. One unit of attention on anticipation is one less unit of attention on experiencing the present. Suppose $a_i(\cdot)$ is the amount of attention on consumption dimension i, and $\sum_{i=1}^{I} a_i(\cdot) \leq A$ is the total amount of attention. Then under certainty, anticipatory utility is given by $Au_p + \sum_i a_i(\cdot)(u_i(z) - u_p)$. The person accurately predicts his preferences and chooses the default, information, and action that maximizes the sum of anticipatory utility and expected consumption utility.

Section 4 presents the results. First, the decision maker will have preferences for information that are a function of the expected consumption utility of the relevant dimension. Keeping the instrumental value of information constant, the decision maker's willingness to pay for information will be increasing in expected consumption utility. Furthermore, we show that if expected consumption utility is sufficiently low, then the decision maker will avoid information even if it would make a consumption utility maximizing agent *strictly better* off for all possible posteriors. An expected consumption utility maximizer and even an anticipating agent with fixed attention (e.g. of the variety found in Eliaz and Spiegler (2006)) both would prefer this type of information. For instance, a patient may avoid information about which procedure would be best for dealing with the patient's heart disease. The less healthy the patient is after the proper procedure, the lower the demand for the information, even though the proper procedure may be very beneficial.

This logic applies analogously to actions. Our second result is that the decision maker will exhibit "behavioral lock-in", in which he may follow the default action even if it fails to maximize expected consumption utility. When the expected consumption utility is low, choosing a non-default action may produce low anticipatory utility. Keeping the expected consumption utility difference between choosing a non-default action and the default constant, the value of choosing the non-default action is increasing in the expected consumption utility of the relevant dimensions.

As an example, we show that a decision maker who has access to partial insurance, but has chosen the default of not purchasing, will continue to not purchase even if new information indicates that the probability of loss is higher than previously believed. An expected consumption maximizer or an anticipating agent with fixed attention would increase insurance. Moreover, we show that if attention shifts sufficiently, and expected consumption utility is sufficiently low, then the demand for insurance is *decreasing* in the probability of loss. As the probability of the bad state increases, thinking about the dimension yields lower anticipatory utility, and hence changing old habits becomes more costly. The model predicts that people will underinsure when the underlying risk increases. A patient who discovers he has heart disease may neglect taking his new heart medication because it shifts attention to his poor health. An expected consumption utility maximizing agent would pay more for insurance when the risk increases.

When information increases attention on a dimension, not only is there a direct effect on anticipatory utility, but also the dimension becomes more heavily weighted relative to other dimensions. In other words, information in our model is a complement to consumption. The information-consumption complementarity can be interpreted as an "advertising effect" in the Becker and Murphy (1993) sense that advertisements are complements to consumption. In an influential paper Becker and Murphy (1993) incorporate advertisement into the traditional "rational" economic framework by modeling an advertisement and a good as complements. Consumption of one increases demand for the the other and advertisements may either directly increase or decrease utility. Our model gives micro-foundations for their model. The complementarity stems from anticipatory utility and shifting attention. Our model also predicts which advertisements will be goods and which will be bads as described in the second result: demand for advertisements will be higher if they provide information about dimensions with higher expected consumption utility.

We show that obtaining information can increase future consumption in a dimension, even when the information implies that marginal consumption utility is low. To illustrate, a patient with heart disease who learns that the benefit to a drug is weaker than expected may still take more of it than if he never received information about the benefit at all. An expected consumption utility maximizer would take less. The reason is that the information itself serves as a reminder of health. For the third result we show that a person who has long had full knowledge of the benefits of an activity (which we model as having a well-informed prior belief) will consume less than a person who has been recently informed about the benefits (which we model as having an uninformed prior belief but a well-informed posterior). The recency of the information increases attention on the dimension and thereby stimulates more consumption in that dimension. Moreover, this implies that two agents with identical preferences and beliefs will behave differently depending on whether information was received recently or in the more distant past. In an economic model without attention this makes no difference. This feature of preferences allows the person to be persuaded by attention-shifting information. A firm will be able to persuade consumers, that is increase their valuation of the product, by providing information that would not otherwise be persuasive to an expected consumption maximizer. This leads to the main application.

In Section 5 we present an application in which a monopolist advertises the exogenous quality of a multi-attribute good to a single consumer. Each attribute corresponds to a consumption dimension of the consumer's utility. The good as a whole is desirable although some attributes may be bads. The standard version of this game is the "Persuasion Game" of Milgrom (1981) and Grossman (1981). In our model, the consumer is additionally influenced by the information-consumption complementarity. The monopolist can increase the price by shifting attention to dimensions in which buying the good yields higher utility than not buying. By the same logic, there will exist an equilibrium where there is no disclosure for attributes in which the utility from the highest quality good is still substantially lower than the utility of not having the good. A drug company may not disclose (without regulation) even "good news" about the side effects of their heart medication because they do not wish to draw attention to these negative attributes. Disclosure will reduce the price that the monopolist can charge.

We then discuss welfare implications and find that the firm's decision to disclose is not affected by how disclosure affects welfare, and so there may be either too much or too little disclosure in equilibrium. The monopolist may shroud an attribute if the gain from separating from lower types is sufficiently less than the consumer's loss in that dimension from buying the good. However, shrouding increases welfare if the utility from the good is greater than the foregone utility. There is no reason to expect that these conditions will have any relationship with each other. Thus the monopolist's decision to shroud will be orthogonal to the welfare effect. We then casually discuss a richer environment in which quality is endogenous. In such an environment, the monopolist may undersupply quality due to an "attentional transaction cost": the monopolist has no means to profitably signal quality and thus has no incentive to produce quality.

In an extension we examine the effect of competition on disclosure using a duopoly model. When the good is sufficiently valuable, the consumers will always buy from one of the two firms. For a given attribute, the firm with higher quality will always do better by disclosing quality. This imposes an "attentional externality" on the second firm who then has no incentive to shroud. The second firm will disclose quality in equilibrium due to the standard unraveling logic. So when the good is sufficiently valuable duopoly results in full disclosure.

In Section 6 we offer potential avenues for future research. We discuss an extended model with multiple periods of anticipation in which the decision maker has incentives to manipulate beliefs for commitment purposes. We also discuss ways to extend the model to analyze risk preferences and hedonic misprediction. Then we conclude.

2 Related Literature

In this section we relate our model to other economic models of anticipatory utility and consider the small economic literature on attention.

Kreps and Porteus (1978) present a revealed preference framework which allows for intrinsic information preferences. In their model a person may have preferences over the timing of the resolution of uncertainty. For example, someone who gets $1 tomorrow if a coin lands heads, may prefer to learn the outcome either now or tomorrow. The authors assume time consistency. They show that within their framework, if a person is indifferent to the timing of the resolution of uncertainty, then preferences could be represented by the standard expected utility model.

Loewenstein (1987) provides the first explicit model of anticipation. In it he assumes that anticipatory utility is time discounted consumption utility. The implication is that people will exhibit what looks like negative time discounting. A person will pay to delay the consumption of goods and will pay to speed the consumption of bads. Loewenstein's model also predicts time inconsistency. A person faced with consuming a good would want to continually delay consumption.

Caplin and Leahy (2001) extend Loewenstein's model to uncertainty. Risk and future expected outcomes map into anticipatory utility. The solution to their model is analogous to subgame perfect equilibrium, and thus, unlike Kreps and Porteus (1978) their model allows for time inconsistency.

This time inconsistency could come from different preference orderings over consumption in the second period, or anticipatory risk preferences that differ from consumption risk preferences. They show that their model is capable of rationalizing the equity premium puzzle and the risk-free rate puzzle.

Kőszegi (2009) extends the work of Caplin and Leahy (2001) to allow expectations to influence utility. Preferences are defined over physical outcomes and beliefs, however unlike Caplin and Leahy (2001) these may interact in the utility function. As a result, a person's expectations of his own actions influences his choice. Since expectations depend on actions, and actions depend on expectations, subgame perfect equilibrium is an insufficient solution concept. Instead, a new solution concept, personal equilibrium, is defined, and extends subgame perfect equilibrium with the added assumption that expectations must be consistent with the actual distribution of outcomes. Kőszegi (2009) shows that if the decision maker does not exhibit informational preferences, multiple equilibria with different utility, or time inconsistency, then preferences could be represented with standard expected utility.

Our model bears similarity with Caplin and Leahy (2001). We begin with a simplified version of their setup, but differ by adding the attention component of anticipation.[1] Preferences that interact attention with anticipatory utility generate behavior that previous models of anticipation cannot capture. Our model relates behavior and information preferences to the *level* of future expected consumption utility. To our knowledge our model is also the first to micro-found an information-consumption complementarity based on anticipatory utility.

Additionally, there is a new growing literature exploring games between agents that have anticipatory utility. Caplin (2002) and Caplin and Eliaz (2003) explore mechanism design with anticipating agents. Kőszegi (2003, 2006), and Caplin and Leahy (2004) model principal-agent behavior when the principal has utility over beliefs. There may be a communication breakdown when a well-intentioned agent interacts with an anticipating principal. Our model continues in this tradition. In Section 5 we specify a game in which a buyer with attention-shifting anticipatory utility may purchase a good from a profit-maximizing seller.

Compared to anticipation, very little economic research has been written on attention. As far as we are aware, there are no papers except for Karlsson et al. (2008), where attention is interpreted as a component of preferences, although there are several economics papers where attention is a constraint or a component of the economic environment. Eliaz and Spiegler (2008) and Masatlioglu et al. (2008) develop models of "limited attention", where the agent is incapable of choosing the optimum over the full choice set, but instead chooses optimally from a subset called a consideration set. Their meaning of attention differs from ours. For them, attention on an option means that the option is

[1] Caplin and Leahy (2001) are the first to discuss the interaction of attention and anticipation, and convincingly motivate it in Section IV.E.

in the consideration set. Attention in our model is purely about anticipating different dimensions of future consumption. Banerjee and Mullainathan (2008) model attention as an input available for the production of two different outputs. One can produce sellable output at their job or domestic output at their home. Gabaix et al. (2006) model attention in a Herbert Simon, satisficing sense. They develop and experimentally test a model of costly information acquisition.

Perhaps the model closest to our own, Karlsson et al. (2008) relates attention to anticipation. The authors present a model of an investor who has reference-dependent utility over money, and an attention function that discretely increases the gain-loss component upon the receipt of information. Our attention function works similarly, but we do not assume reference-dependent utility. They predict information-seeking when the market is going up and information avoidance when the market is going down. Our model makes similar predictions, but ours are based on the absolute level of consumption utility and not the change. Furthermore, our model relates actions to attention, and we predict that people will follow their default when expected consumption utility is low. We also have multiple dimensions of consumption which allows for the information-consumption complementarity.

Finally, there is a family of mostly recent papers that interacts anticipatory utility or other belief-based preferences with belief distortion. The basic concept in these models is that the decision maker gets utility directly from his beliefs and has a technology to distort his beliefs. The tradeoff in these models is between happy beliefs and accurate decisions. Starting with Akerlof and Dickens (1982), workers may choose to be fearful and appropriately use safety equipment, or blissfully ignore dangers while facing bodily harm. Brunnermeier and Parker (2005) generalize and apply this to portfolio choice and consumption-savings decisions. They show that investors will overestimate returns and consumers will overconsume. Bracha and Brown (2007) apply a similar model to insurance and show that with copayments, the worse a bad outcome is, the less insurance the agent will purchase. In Mayraz (2009), beliefs are not chosen but distorted as a function of payoffs. Unlike these models, our relevant tradeoff is not between accurate beliefs and happy beliefs, but rather consumption utility and the attention-weighted consumption utility (anticipatory utility). The driving force in the belief distortion models is the utility difference between mutually exclusive events. The driving force in our model is the difference between the expected consumption utility of a dimension and the forgone utility of attention. Since consumption in dimensions is not mutually exclusive, there is a clear substantive distinction between dimensions and events. Because of this, we address different questions. Our model directly considers informational preferences, which is an unexplored topic in the belief-distortion literature, and novelly predicts that demand for information is increasing with expected consumption utility, and that furthermore there is an information-consumption complementarity.

Psychological evidence for the assumptions of our model is desirable. There are a small handful of psychology experiments that are related (Averill and Rosenn, 1972; Miller and Mangan, 1983; Wood et al., 2002), but they do not specifically test the psychological behaviors that we assume as the basis for our model. An additional benefit to specifying a formal model is to help experimentalists identify the role of anticipatory utility in behavior.

3 Model

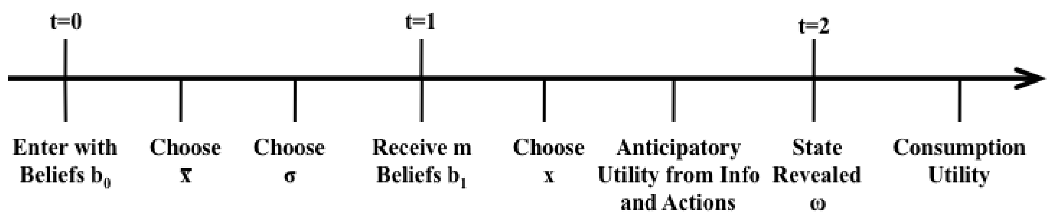

Figure 1: Timeline

In this section we present the formal model. We consider a central focus of the model to be the specification of the dependency of attention on actions and information. The timing is illustrated in Figure 1. There are three time periods, $t = 0, 1, 2$. At $t = 0$ the decision maker (henceforth referred to as the "DM") chooses a default and information, at $t = 1$ the DM chooses an action and experiences anticipatory utility, and at $t = 2$ the DM experiences consumption utility.

Beginning with the economic environment, let the action set X be a nonempty finite set. Initially at the beginning of $t = 0$, the DM chooses a default action $\bar{x} \in X$ to follow. A default is an action, not a strategy. It is an action that is specified in advance and cannot be made contingent on future information gathering or changes in beliefs.

The purpose of the default is to specify the action for which no attention is required. In the model, both a change in actions and a change in beliefs shift attention. Whereas the natural default belief is the prior, there does not seem to be a similarly natural default for actions. Thus as a starting point, we assume that the DM chooses his default, although in general we interpret the default action as the activity or inactivity that would naturally occur without the exertion of attention, and this need not correspond to the utility-maximizing default.

We have several interpretations of the default. The default may represent habits, automatized or routinized actions, or a plan of action: washing one's hands after using the restroom, choosing one's usual dish instead of considering the full menu, or a plan to work eight hours this week on a project. In other cases, the default is most appropriately the action that best corresponds to omission. Leaving one's stock portfolio as is, an act of omission, would appropriately be the default.[2] We assume that the default is a non-contingent plan because it seems psychologically likely that following a full belief-contingent plan requires attention. If one discovers it necessary to eliminate salt from one's diet for health reasons, attention is required to accomplish this, and doing so will remind oneself of one's health. If the default were a full belief-contingent strategy, our model would have no new predictions regarding the role of actions and anticipation. The DM would be able to specify a default that he could follow in equilibrium for all beliefs, and attention would never shift through actions. In some cases our assumption that the DM may choose the default may be implausible. In this paper we will try to be clear on how the results depend on the default, and in many parts of this paper the results will not depend on the assumption that the default is ex ante optimal.

Let Ω be the non-empty finite state space. A state $\omega \in \Omega$ represents the relevant exogenous characteristics of the world that affect the physical consequence. For example, states may represent the weather in Hawaii or one's genetic proclivity to getting heart disease. Let $\Delta(\Omega)$ represent the set of all probability distributions on the set Ω. Define $b_1 \in \Delta(\Omega)$ as the beliefs at the end of period $t = 0$, and $b_0 \in \Delta(\Omega)$ is the prior.

Let Z be the non-empty finite set of possible physical consequences in period 2. These represent all the outcomes that the DM cares about that can occur in the future. For example, a consequence $z \in Z$, may be having an enjoyable vacation next month, and a healthy heart ten years into the future. Another consequence may be having a stressful vacation, and heart disease ten years into the future. The default has no direct influence on physical consequences z. In $t = 1$, after receiving information, the DM chooses an action $x \in X$ that affects z. The action $x \in X$ is chosen from the same action set as the default $\bar{x}$ and need not be the same. Actions map from states to consequences, $x : \Omega \to Z$. An action is a lottery over consequences.

At $t = 1$, the DM may receive information. Let S be a finite set of realized signals. A signal is a function that maps from the state to a probability distribution over realized signals $\sigma : \Omega \to \Delta(S)$. To give an example, the states of the world would be various genetic proclivities to heart disease, getting a genetic test would be a signal, and learning that one has a gene that makes one prone to heart disease would be the realized signal. At $t = 0$ the DM may choose a signal from the set of signals $\sigma \in \Sigma$. Beliefs are updated in a Bayesian way $b_1 : \Sigma \times S \to \Delta(\Omega)$.

Now we discuss preferences. Consumption utility is defined over physical consequences. Consump-

[2]It is unclear what an act of omission means in a formal sense. There does not exist a formal definition, to our knowledge, in game theory or decision theory. Given an action set, we interpret the act of omission to be the act that would be "chosen" by a person who is incapacitated or asleep. Specifying the omitted action then requires contextual information beyond the traditionally defined economic environment.

tion utility is additively separable in I dimensions. We offer two interpretations of these dimensions. First, they may represent different time periods. For example, one dimension may represent consumption of entertainment tomorrow versus consumption of entertainment 21 days from now. There is a tradition in economics of utility functions that are additively separable over time periods. If the dimensions represent non-contemporaneous consumption episodes, then the additive separability is consistent with standard economic models. Here the periods need not represent time periods of equal duration but rather periods that are psychologically contiguous. Under this interpretation, period 2 is composed of I sequential consumption episodes. Second, a dimension may represent contemporaneous consumption of psychologically distinct aspects of an episode. For example, tomorrow one may derive consumption from both lunch and from the clothes that one wears. These two forms of consumption may overlap in time, but as long as the taste of the food is independent of the comfort from wearing the clothes and vice versa, we can separate these forms of consumption in the utility function. Under this interpretation, I forms of consumption occur contemporaneously over the duration of period 2. In either interpretation the critical psychological assumption is that these dimensions represent forms of consumption that could be anticipated independently of each other. Finally, consumption utility in dimension i is given by the vNM utility function $u_i : Z \to \mathbb{R}$ and total consumption utility is $\sum_i^I u_i(z)$. Since utility is defined over consequences, it is convenient to define an expected consumption utility function over actions. Let $w_i(x; b_t) \equiv \sum_{\omega \in \Omega} \Pr_{b_t}(\omega) u_i(x(\omega))$.

Anticipatory utility is experienced at the end of $t = 1$. We assume that anticipatory utility is a weighted sum of expected consumption utility. Each dimension is weighted by the attention on that dimension. Let $a_i(\bar{x}, x; b_0, b_1)$ be the attention on dimension i, which depends on the default action, the realized signal, and the action. The anticipatory utility experienced in dimension i is $a_i(\bar{x}, x; b_0, b_1) w_i(x; b_1)$.

We assume that there is a fixed amount of attention meaning that if one is not thinking about the future then one is thinking about something else (such as the Johnny Depp movie playing on TV), and deriving a flow of utility from that alternative. A critical feature of the model is that there is a tradeoff between attending to future consumption and attending to the present. Fixing the total amount of attention is the means by which we impose these tradeoffs. For simplicity we assume that there is a constant forgone utility to attention. For every unit of attention on the future, there is one less unit of attention on an experience with utility level u_p, which may be interpreted as the utility from the present. The present utility $u_p \in \mathbb{R}$ is exogenous. This assumption serves a role analogous to quasi-linear utility for money. Quasi-linear utility for money has a constant forgone utility to consumption, in our model there is a constant forgone utility to attention. Assuming a constant forgone utility is a simplifying assumption that makes the model more tractable.[3] A richer model might have the foregone utility of attention as a function of current consumption (i.e. $u_p : X \to \mathbb{R}$). Total anticipatory utility is given by $\sum_{i=1}^{I} a_i(\bar{x}, x; b_0, b_1)(w_i(x; b_1) - u_p)$.[4]

Now we define the attention function. We make two critical assumptions. First, we assume that information shifts attention to dimensions that have a different ex post expected consumption utility than ex ante expected consumption utility. When one learns about next month's forecasted weather in Hawaii, one thinks about next month's vacation, and not about one's health. We formalize this by assuming that if $w_i(x'; b_1) \neq w_i(x'; b_0)$ for some $x' \in X$, then $\gamma > 0$ attention shifts to dimension i. In other words, receiving a signal that changes expected consumption in i for at least one action, discretely shifts γ units of attention to i. In some ways this assumption is pretty strong. If a person receives information about an option x that changes expected consumption utility in dimension i, attention will shift even if x would never be chosen under any beliefs. For example, if the DM receives information about a vacation to Antarctica, even if the DM would never choose to go there, the DM anticipates his expected vacation to Hawaii more. However, this is somewhat more plausible than assuming that attention shifts only if the information affected utility given the chosen action. Under such an assumption, when considering a vacation to Hawaii, receiving information about Maui will not cause attention to shift to vacation if the DM ultimately decides instead to go to O'ahu.

The second critical assumption is that choosing a non-default action shifts attention to dimensions that have a different consumption utility than with the default action. The interpretation is that deviating from habits, routines, or plans requires attention, but when following the default the DM need not think about the consequences while in the moment. If one is in the habit of eating fatty food and not taking heart medication, then changing one's diet and taking medication makes one think more of one's health.

We model this by assuming if $w_i(x; b_1) \neq w_i(\bar{x}; b_1)$ or if $w_i(x; b_1) \geq u_p$, then $\alpha > 0$ attention shifts to dimension i. In other words, choosing an action that deviates from expected default consumption in i discretely shifts α attention to i. Additionally, whenever expected consumption utility is above the forgone utility, the DM will think about consumption. This second condition indicates that the DM will attend to any dimension when it is pleasurable to do so.[5] The interpretation is that the DM can

[3] An alternative assumption would be to fix all the attention weights $\sum a_i(\cdot) = 1$. The basic forces in such a model would be the same. This approach has an additional complication because increasing attention in dimension i would necessarily decrease attention in other dimensions. We would then need to assume which dimensions decrease in attention and by how much. The "constant foregone utility" assumption is simpler in this respect.

[4] For clarity, as mentioned in the introduction, one can think of total attention as equal to $A \geq \sum_{i=1}^{I} a_i(\bar{x}, x; b_0, b_1)$. Then anticipatory utility is equal to $Au_p + \sum_{i=1}^{I} a_i(\bar{x}, x; b_0, b_1)(w_i(x; b_1) - u_p)$. If we subtract the constant Au_p we get the desired expression. Subtracting this constant will have no effect on the preference ordering.

[5] This second condition eliminates burning money problems. Without it, the DM may "burn utils" in high-level dimensions to shift attention to that dimension. For example, suppose the DM may choose between two vacations of differing quality and

freely pay attention and will pay attention to pleasurable future experiences regardless of the default action. For purely expositional purposes, until we reach Section 5, we will assume that the anticipatory utility is always less than the forgone utility of attention, $w_i(x; b_t) < u_p$ for all i, x, and b_t. When $w_i(x; b_t) \geq u_p$ the default action becomes irrelevant to the DM's choice, so in this sense, behavior will be closer to an expected consumption utility maximizer.

The two critical assumptions are meant to be analogous. Deviations from a default shift attention to consumption that is affected by the deviation. In terms of information, we explicitly assume that the default beliefs are the prior. Shifting attention by way of information may be a different psychological process than shifting attention by way of actions. Both can be incorporated into our reduced form framework in a similar way, although we do not claim that the psychological process is identical. To incorporate new information, one typically needs time to think about the implications of the information. Likewise, we speculate that when one takes a non-default action, it requires time to think about the consequences of the action.[6]

The total shift in attention comes from the attention shifted from information plus the attention shifted from actions. We interpret both of these as increasing the amount of time spent anticipating future consumption. Taking the view of Kahneman (2000) in his moment-based approach to experienced utility, more time anticipating a future episode leads to more total anticipatory utility from that episode. The attention function is given below.

Attention Function:

$$\begin{aligned} a_i(\bar{x}, x; b_0, b_1) \equiv &\gamma \mathbf{1}\{w_i(x'; b_1) \neq w_i(x'; b_0) \text{ for some } x' \in X\} \\ &+ \alpha \mathbf{1}\{w_i(x; b_1) \neq w_i(\bar{x}; b_1) \text{ or } w_i(x; b_1) \geq u_p\} \end{aligned} \tag{1}$$

We make attention discrete for two reasons. First, we do so because it is simple and we avoid issues regarding the curvature of attention. Understanding the effects of the curvature of the attention

suppose the default is the high-quality destination. Without the second condition, the DM may choose to go to the inferior vacation for the sole purpose of shifting attention to the vacation dimension. More generally, if the DM's only means to deviate a high-level dimension's utility from the default utility is to lower it, then the DM will frequently take this option to shift attention. This seems psychologically implausible. It seems psychologically more plausible that a person can willingly shift one's attention to high-level dimensions. If one wants to spend more time thinking about one's vacation, one does not need to make the vacation worse in order to think about it more. Furthermore, without this assumption the DM has an incentive to choose defaults that he has no intention of following for the sole reason of shifting attention to the dimension.

Fundamentally, this builds an asymmetry into the model. People can willingly think about high-level dimensions independent of their actions, but cannot avoid thinking about low-level dimensions independent of their actions.

[6] When choosing a non-default action, one also requires more time to think about which action is optimal. One directs attention to a dimension in order to optimize that dimension. This interpretation, although in the spirit of the model, differs somewhat from how the model works. Suppose, $\bar{x} = x'$ and choosing $x = x''$ gives a different expected consumption utility in both the vacation and health dimensions. According to the model, choosing x'' would shift attention to both dimensions. However, in the directed attention interpretation, if the person chooses x'' to optimize the vacation dimension, attention will shift only to vacation even though both vacation and health are affected by the choice.

function may be a useful direction for future research. Second, in some situations discreteness is plausible. For example, if one is shopping for a camera and if "not buying" is the default, one spends a chunk of time thinking about future episodes in which one uses the camera. Our discreteness assumption implies that the amount of time thinking about the camera is independent of which camera is purchased as long as some camera is purchased. However, in some other cases, this discreteness assumption would appear to be inappropriate. It is quite likely that the more time a worker puts into a project, the more attention the worker has on the consequence of the project. Although, our model is unable to capture this effect, it would capture attention shifting when the chosen work level differs from the default work level.

We do not believe that our model exhausts all the factors that affect attention. Proximity, framing, vividness, and other psychological properties of stimuli certainly play a role. However, we focus on actions and information for two reasons. First, we think that actions and information have a strong effect on attention in important ways. Second, actions and information are important economic concepts, so exploring how they affect attention and therefore economic decision making may yield important insights into economic behavior.

The timing is illustrated in Figure 1. There are three choices that are made sequentially. Going in order, at $t = 0$ the DM chooses a default $\bar{x} \in X$, and then chooses information $\sigma \in \Sigma$. At $t = 1$, information arrives and then the DM chooses x. The objective function for each period is:[7]

$$V_0(\bar{x}, \sigma; x, b_0) \equiv E_{b_0}\left[\sum_{i=1}^{I}(1 + a_i(\bar{x}, x; b_0, b_1))w_i(x; b_1) - a_i(\bar{x}, x; b_0, b_1)u_p\right] \tag{2}$$

$$V_1(x; \bar{x}, b_0, b_1) \equiv \sum_{i=1}^{I}(1 + a_i(\bar{x}, x; b_0, b_1))w_i(x; b_1) - a_i(\bar{x}, x; b_0, b_1)u_p \tag{3}$$

Notice that if $\gamma = \alpha = 0$, the DM is an expected consumption utility maximizing agent. With $\gamma = \alpha = 0$, attention on all dimensions i is always $a_i(\bar{x}, x; b_0, b_1) = 0$. The objective function then becomes $V_1(x; \bar{x}, b_0, b_1) = \sum_{i=1}^{I} w_i(x; b_1)$, which is the objective function of an expected consumption utility maximizer with I additively separable dimensions. Thus the standard model is nested within our model. Throughout the paper we will compare an agent with $\gamma = \alpha = 0$ to an agent who has $\gamma, \alpha > 0$.

[7]Preferences in this model are time inconsistent as they are in other models of anticipatory utility (Loewenstein, 1987; Caplin and Leahy, 2001; Kőszegi, 2009). Total expected utility at the beginning of period two before the resolution of uncertainty is given by $V_2 \equiv \sum_{i=1}^{I} w_i(x; b_1)$. The time inconsistency arises because the $t = 0$ and $t = 1$ selves have different preferences than the $t = 2$ self for $t = 2$ consumption. For our simple setup this does not matter, since no decisions are made at $t = 2$. If we were to have actions at $t = 2$ the analogous solution concept for the model would be subgame perfect equilibrium. A similar structure is used in Caplin and Leahy (2001). Their model is two periods with time inconsistent preferences, and their solution concept is analogous to subgame perfect equilibrium.

With specified objective functions we can now define equilibrium. Since actions will be a function of beliefs, we must define strategies. A period 1 *strategy* $\psi(b_0, b_1)$ is a function that assigns a default action $\bar{x} \in X$ and a signal $\sigma \in \Sigma$ to the first two decision nodes respectively, and an action $x \in X$ for every subsequent information set. We denote a strategy with the triple $(\bar{x}(b_0), \sigma(b_0), x(b_0, b_1))$. The model is solved backwards. The optimal action is $x^*(b_0, b_1) \equiv arg\,max_{x \in X} V_1(x; \bar{x}, b_0, b_1)$. The optimal signal is $\sigma^*(b_0) \equiv arg\,max_{\sigma \in \Sigma} V_0(\bar{x}, \sigma; x, b_0)$. And the optimal default action is $\bar{x}^*(b_0) \equiv arg\,max_{\bar{x} \in X} V_0(\bar{x}, \sigma; x, b_0)$.

Definition The *optimal strategy* is defined as $\psi^*(b_0, b_1) = (\bar{x}^*(b_0), \sigma^*(b_0), x^*(b_0, b_1))$.

When we take a decision problem in which the DM receives no new information the DM behaves just like a consumption-maximizing agent.

Proposition 1 *If for all $\sigma \in \Sigma$, $b_1 = b_0$ for all realizations of the signal, then $x^*(b_0, b_1)$ maximizes expected consumption utility.*

When beliefs do not change, the DM will choose a consumption utility maximizing default and follow it through. Since we assumed $w_i(x; b_t) < u_p$ for all i and x, it is optimal to choose $\bar{x} = x$ otherwise attention will shift to low-level dimensions. Since all dimensions are weighted equally, the consumption utility maximizing bundle will maximize total utility.[8] More simply put, when the DM does not receive any new information and expected consumption utility is low in all dimensions, the DM maximizes expected consumption utility. No attention will be on future consumption. New information will be the key component of the environment that induces novel behavior.

When there is new information, but some dimensions have high expected consumption utility and some have low, the DM will have α attention shifted to only the high dimensions. Since these dimensions are weighted more heavily, the DM will consume more in these dimensions than a consumption utility maximizer would. To an observer, when there is no action in $t = 2$, the DM's behavior would be indistinguishable from a consumption maximizer. However, if there were actions in $t = 2$ then the DM may behave in a time inconsistent manner (i.e. the DM may pay for commitment in the absence of uncertainty).

[8]If $w_i(x; b_1) \geq u_p$, or for all i and x, the proposition would remain true. If all the dimensions are greater than u_p, then α attention will shift to all of those dimensions and so dimensions will be weighted equally. It follows that the consumption utility maximizing bundle will maximize total utility.

4 Results

4.1 The Level of Expected Consumption Utility

The DM prefers to pay attention to dimensions that have high expected consumption utility. Taking a non-habitual action such as planning for one's vacation is far more enjoyable than doing similar planning tasks for a dentist's appointment. An agent with $\gamma = \alpha = 0$ consumes as long as the marginal benefit to consumption utility exceeds the marginal cost. Our attention-sensitive anticipatory agent cares both about the marginal benefit of consumption and the *level* of consumption utility. Lower level dimensions offer a less pleasurable anticipatory experience. It follows that the DM is incentivized to shift attention away from low-level dimensions through the information and action that he chooses.

Information

An agent with $\gamma = \alpha = 0$, without strategic concerns would never avoid information. Our model generates intrinsic information preferences that are a function of the level of expected consumption utility. Receiving information about a dimension increases attention and hence anticipatory utility from that dimension. If the expected consumption utility in a dimension is sufficiently low, the DM will prefer to avoid information that is relevant to the dimension. Unlike some other models with anticipatory utility, information avoidance may occur in our model even when information is purely "how" information; as in "how to choose the right action". When we use the term "how" information, we refer to a signal that with any realization would increase consumption utility by a fixed amount given consumption-maximizing behavior.

Definition Let $x^c(b_t) \equiv arg\, max_{x \in X} \sum_{i=1}^{I} w_i(x; b_t)$. A signal σ contains *purely "how" information* if for all realizations of the signal s in the support of σ give, $\sum_{i=1}^{I} w_i(x^c(b_1); b_1(s)) > \sum_{i=1}^{I} w_i(x^c(b_0); b_0)$ and for any two realizations of the signal in the support $s, s' \in S$, $\sum_{i=1}^{I} w_i(x^c(b_1); b_1(s)) = \sum_{i=1}^{I} w_i(x^c(b_1); b_1(s'))$.

After receiving a purely "how" signal, a consumption utility maximizer will have a fixed higher utility regardless of the realization of the signal. In other words, before receiving the signal, the DM knows exactly what her expected consumption utility could be. An example of purely "how" information would be receiving instructions on how to bake grandma's well-known cake. A person with the proper ingredients and the wrong instructions will make a mess, but a person with grandma's recipe will bake a good cake of known quality. The person's utility is independent of the procedure, as long as following the instructions produces an instance of grandma's cake.

In other models with intrinsic information preferences (Kreps and Porteus, 1978; Kőszegi, 2003; Caplin and Leahy, 2004; Eliaz and Spiegler, 2006), the agent always prefers purely "how" information.

In several models (Kőszegi, 2003; Caplin and Leahy, 2004), anticipatory utility is either a concave or convex function of expected consumption utility. Information that has no instrumental benefit adds a random variable with mean zero to expected utility. As a result, anticipatory utility that is concave in expected consumption utility produces information-averse preferences, and anticipatory utility that is convex in expected consumption utility produces information-seeking preferences. To the contrary, purely "how" information is purely instrumental. Ex ante, obtaining purely "how" information yields a known expected consumption utility. In other words, expected consumption utility becomes a degenerate random variable. With purely "how" information, information preferences are not relevant, and since the frontier of expected consumption utility increases with purely "how" information, the agents in these models will always prefer the information.

Example : Avoiding Purely "How" Information

In this example we show that an anticipating agent with attention will avoid purely "how" information when expected consumption utility is low.

The DM has two actions available to him $X = \{j, k\}$. The state denotes which option is better, $\Omega = \{j, k\}$.[9] The action and the state affect only dimension i. The DM has access to two signals, one that reveals the state, and one that provides no information, $\Sigma = \{r, n\}$. Notice that the signal r contains purely "how" information. The optimal default $\bar{x}^*$ will depend on which signal is chosen, but the main conclusion will be true for any default.

Total utility from the two signals is given by:

$$\begin{aligned}
V_0(\bar{x}^*(b_0), n; x^*(b_0, b_0), b_0) = & E_{b_0}\Big[\big(1 + a_i\big(\bar{x}(b_0), x^*(b_0, b_0); b_0, b_0\big)\big)\, w_i(x^*(b_0, b_0); b_0\big) \\
& - a_i\big(\bar{x}(b_0), x^*(b_0, b_0); b_0, b_0\big) u_p\Big] \\
V_0(\bar{x}^*(b_0), r; x^*(b_0, b_1), b_0) = & E_{b_0}\Big[\big(1 + a_i\big(\bar{x}(b_0), x^*(b_0, b_1); b_0, b_1\big)\big)\, w_i(x^*(b_0, b_1); b_1) \\
& - a_i(\bar{x}(b_0), x^*(b_0, b_0); b_0, b_0) u_p\Big]
\end{aligned}$$

This can be simplified because in the absence of new information, the optimal default and optimal action will be the same and so attention will not shift to dimension i, $a_i\big(\bar{x}(b_0), x^*(b_0, b_0); b_0, b_0\big) = 0$. We wish to find conditions on the parameters such that no information is preferred $V_0(\bar{x}^*(b_0), n; x^*(b_0, b_0), b_0) > V_0(\bar{x}^*(b_0), r; x^*(b_0, b_1), b_0)$, which implies

$$u_p > \frac{E_{b_0}\left[\big(1 + a_i\big(\bar{x}(b_0), x^*(b_0, b_1); b_0, b_1\big)\big) w_i(x^*(b_0, b_1); b_1)\right] - w_i(x^*(b_0, b_0); b_0)}{E_{b_0}[a_i\big(\bar{x}(b_0), x^*(b_0, b_1); b_0, b_1\big)]}.$$

[9] There are two consequences, matching the state and mismatching, $Z = \{ma, mi\}$. Matching is preferred $u_i(ma) > u_i(mi)$.

The denominator of the right-hand side term $E_{b_0}\left[a_i\left(\bar{x}(b_0), x^*(b_0, b_1); b_0, b_1\right)\right] \geq \gamma$ since information changes expected consumption, which makes this term positive. This means that if u_p is relatively large compared to the consumption utility, it will never be optimal to receive information. Receiving the information shifts attention to dimension i which is undesirable to think about. An agent with $\gamma = 0$ and even an agent with "standard" anticipatory utility with fixed attention would always choose $\sigma = r$.

As long as utility from the dimension is sufficiently low, the DM will avoid information. In fact, as the utility from the dimension decreases, the total benefit from information goes down. We can see this looking at the numerator of the expression above. Keeping the instrumental benefit of information constant $w_i(x^*(b_0, b_1); b_0) - w_i(x^*(b_0, b_0); b_0)$, reducing the level of both $w_i(x^*(b_0, b_1); b_0)$ and $w_i(x^*(b_0, b_0); b_0)$ decreases the numerator. This increases the range of u_p for which no information is preferred.

Information avoidance plausibly occurs in various important facets of life. Our prediction is that it will occur more when expected consumption utility is low. Health, may be a particularly dread-provoking topic. Upon finding early evidence of a breast tumor many women delay seeking help (Nosarti et al., 2000; Bish et al., 2005). Rather than getting a diagnosis and treatment, many women avoid confronting the issue until the situation is substantially worse. Lindberg and Wellisch (2001) find that breast self exams are negatively correlated with anxiety. Smith and Croyle (1995) find that 16.1% of their sample would not want to get a genetic test for susceptibility to colon cancer. Many at risk individuals avoid STD tests, despite the benefits of diagnosis and treatment. For example, more than half of those diagnosed with AIDS between January 1990 and December 1992 first tested for HIV within one year of their AIDS diagnosis (Chesney and Smith, 1999), even though the onset of AIDS typically occurs a decade after HIV infection.[10] This suggests that many patients had HIV for nearly a decade before testing despite the benefits of diagnosis.

The proposition below formalizes this logic.

Proposition 2 *Let σ induce beliefs b_1 such that there is a positive probability that $w_j(x; b_1) \neq w_j(x; b_0)$ for some x, and σ_n induces beliefs $b_1 = b_0$ for all realizations of the signal. Then keeping $E_{b_0}[w_i(x; b_1)] - w_i(x; b_0)$ constant for all i, the difference $V_0(\bar{x}, \sigma; x^*(b_0, b_1), b_0) - V_0(\bar{x}, \sigma_n; x^*(b_0, b_0), b_0)$ is strictly increasing in $E_{b_0}[w_j(x^*(b_0, b_1); b_1)]$.*

[10] According to the Center for Disease Control: "Before antiretroviral therapy became available in 1996, scientists estimated that AIDS would develop within 10 years in about half the people with HIV." http://www.cdc.gov/hiv/resources/brochures/livingwithhiv.htm#q3

The proposition says that increasing the expected consumption utility of a dimension that is affected by the information, while keeping the instrumental value of information constant, increases the willingness to pay for the information. The reasoning is straightforward. The signal shifts γ attention to the dimension. Information that shifts attention to high-level dimensions is more desirable.

Corollary 1 *Suppose the DM has a degenerate action set $X = \{x\}$. Let σ induce beliefs b_1 such that there is a positive probability that $w_j(x; b_1) \neq w_j(x; b_0)$. Then the difference $V_0(\bar{x}, \sigma; x^*(b_0, b_1), b_0) - V_0(\bar{x}, \sigma_n; x^*(b_0, b_1), b_0)$ is strictly increasing in $w_j(x; b_0)$.*

The corollary simply restates Proposition 2 when the action set is degenerate. If there is no choice of action, then information that changes expected consumption utility in only dimension j is more desirable the higher the expected consumption utility of dimension j. This formulation can be used to identify the expected consumption utility of a dimension. If the dimensions of consumption are known, then a person's information preferences for a dimension will correspond to the expected consumption utility of that dimension. The preference ordering of signals that are only informative about a single dimension, will correspond with the same ordering of expected consumption utility levels. Thus it is possible, under these special conditions, for an observer to infer relative expected consumption utility levels.

Actions

In this section we show that the DM will exhibit "behavioral lock-in" by not fully adjusting to new information when the expected consumption utility is low. The DM wishes to follow the default when consumption utility is low to avoid negative anticipatory feelings. A patient, recently diagnosed with heart disease, is advised to stick to a strict low-fat diet. Adjusting habits are naturally costly. But adjustment is even more challenging when the new diet is a constant reminder of one's illness. The important factor for behavioral lock-in is that the optimal deviation still produces a low consumption utility. If the new diet could completely cure heart disease, then the individual will adjust.

The model predicts that as the achievable expected consumption utility declines in a dimension, ceteris paribus, the DM has a stronger preference to choose the default. If expected consumption utility in a dimension becomes sufficiently low, the anticipatory cost to deviating from the default may exceed the benefit to consumption, and as a consequence the DM will choose the default. We illustrate this in an example below.

Example : Partial Insurance

In this example the DM faces a risk in a low-level dimension. We show that under plausible conditions, the DM will avoid partial insurance. Moreover, the DM's willingness to pay for insurance will be *decreasing* in the probability of the low state.

The action set has two elements, partial insurance and no insurance, $X = \{p, n\}$. The partial insurance could be a precautionary action that reduces the risk of a bad consequence, like taking medication to reduce the risk of heart attack, or using protection to reduce the risk of getting an STD. There are two states, high and low, $\Omega = \{h, l\}$. Only dimension i is affected by this choice and state. The DM has the prior that the low state occurs with probability $\Pr_{b_0}(l) \in (0, 1)$. The choice set for information is degenerate with only one signal $\Sigma = \{\sigma\}$ which provides information about the probability of the low state. Given the prior, there will be an optimal $\bar{x}$. We will consider both $\bar{x}^* = p$ and $\bar{x}^* = n$.

This is a reduced form approach to insurance. Let $\hat{b}$ be the beliefs at which the DM is indifferent between the two actions $w_i(p; \hat{b})) = w_i(n; \hat{b})$. Furthermore, expected consumption utility is decreasing in the probability of the low state $\frac{\partial}{\partial \Pr_{b_t}(l)} w_i(x; b_t) < 0$, but the relative benefit of insurance is increasing in the probability bad state $\frac{\partial}{\partial \Pr_{b_t}(l)} (w_i(p; b_t) - w_i(n; b_t)) > 0$. The timing goes: the DM chooses a default $\bar{x} \in X$, the DM receives σ which updates the probability of the low state to $\Pr_{b_1}(l)$, the DM chooses $x \in X$, experiences anticipatory utility, and then consumption utility.

We solve backwards. First, suppose the default is $\bar{x} = p$. Then

$$V_1(p; p, b_0, b_1) = (1 + \gamma) w_i(p; b_1) - \gamma u_p$$

$$V_1(n; p, b_0, b_1) = (1 + \alpha + \gamma) w_i(n; b_1) - (\alpha + \gamma) u_p$$

Here we can see that if $b_1 = \hat{b}$, then $V_1(p; p, b_0, b_1) > V_1(n; p, b_0, b_1)$. So when the default is partial insurance, partial insurance will be chosen for lower levels of $\Pr_{b_1}(l)$ relative to an expected consumption utility maximizer. Deviating from the default of buying insurance induces low anticipatory utility because the attention shifts to a dimension with low expected consumption utility. This additional cost incentivizes the DM to stay with the default.

Now suppose the DM chooses $\bar{x} = n$, then the utility of the two options are

$$V_1(p; n, b_0, b_1) = (1 + \alpha + \gamma) w_i(p; b_1) - (\alpha + \gamma) u_p$$

$$V_1(n; n, b_0, b_1) = (1 + \gamma) w_i(n; b_1) - \gamma u_p.$$

Here we can see that when $b_1 = \hat{b}$, then $V_1(n; n, b_0, b_1) > V_1(p; n, b_0, b_1)$. So when the default is no insurance, no insurance will be chosen for lower levels of $\Pr_{b_1}(l)$ relative to what an expected consumption utility maximizer would choose. To summarize, when the affected dimension has low expected consumption utility, and a large amount of attention shifts from actions, the DM will stick to old plans more often than an agent with $\gamma = \alpha = 0$ would. If he initially plans no insurance, he will get no insurance more often than a $\gamma = \alpha = 0$ agent, and if he initially plans to get insurance then he will get insurance more often than a $\gamma = \alpha = 0$ agent.

Furthermore, under a plausible condition the utility difference between choosing p versus n when the default is n, $V_1(p; n, b_0, b_1) - V_1(n; n, b_0, b_1)$, may *decrease* with $\Pr_{b_1}(l)$. If

$$\alpha > \frac{(1+\gamma)\frac{\partial}{\partial \Pr_{b_1}(l)}(w_i(p; b_1) - w_i(n; b_1))}{-\frac{\partial}{\partial \Pr_{b_1}(l)} w_i(p; b_1)} \Rightarrow \frac{\partial}{\partial \Pr_{b_1}(l)}\big(V_1(p; n, b_0, b_1) - V_1(n; n, b_0, b_1)\big) < 0$$

This expression says that when α is sufficiently large, the DM will pay *less* for insurance when the probability of the bad state *increases*. The explanation is that the anticipatory cost of deviating from the plan is increasing with the probability of the bad state. When anticipatory utility is large, this effect exceeds the consumption benefit of becoming partially insured.[11]

Will the DM ever choose $\bar{x} = n$ in equilibrium? If $\Pr_{b_0}(l)$ is sufficiently low, it may be optimal to set $\bar{x} = n$. If the DM is surprised by learning that the probability of the low state is very high, the DM may continue to avoid insurance and will actually pay more to avoid it as the bad state becomes more likely.[12]

What would happen if insurance were full? It depends on whether the utility in the good state is high $u_i(x(\omega)) \geq u_p$. If so, then with full insurance there would be no associated anticipatory cost to deviating from the default. A higher probability of the bad state could only increase the willingness to pay for insurance, just like for an agent with $\gamma = \alpha = 0$. However, if full insurance still left consumption utility low $u_i(x(\omega)) < u_p$, then even if p provided full insurance, the analysis above would apply.

We speculate that the logic presented in the example above may be a cause for a variety of real-world situations in which people underinsure. A person with heart disease may take medication which lowers the risk of a heart attack. However, using the medication may be a constant reminder of one's illness. Compliance rates with physician-recommended health regimens are often quite low (Becker and Maiman, 1975; Sherbourne et al., 1992; Lindberg and Wellisch, 2001; DiMatteo et al., 2007). Greater

[11]We should note that this comparative static is driven by the fact that the person does not want to pay attention to the decision. If for some reason the person's attention were shifted to the dimension (i.e. there are multiple levels of coverage and the person is required to get some insurance), then demand for insurance would be increasing with risk as standard theory would predict.

[12]There are three forces at work here. The first force is the one emphasized: the DM does not want to deviate in dimensions with low expected consumption utility. The second force is the information-consumption complementarity described in greater detail later in the paper. Receiving information about the risk shifts attention to i which gives the DM more incentive to buy insurance. Depending on the relative magnitude of α versus γ, receiving the information could either induce greater demand for insurance or not. The third force is the standard incentive — the DM wants to maximize consumption utility. Either way, the cost of deviating is decreasing in the expected consumption utility under the action to which the DM deviates.

fear, lower expectation of health outcome, and lower self-reported health has been correlated with less adherence (Sherbourne et al., 1992; DiMatteo et al., 2007). Moreover, in more serious illnesses such as cancer, HIV, and heart failure those with objectively poorer health are *less* likely to be adherent (DiMatteo et al., 2007).

Condom use can be considered a form of partial insurance against sexually transmitted diseases. The standard direction of causation is that less condom use will result in greater HIV transmission. However, our model suggests that reverse causation may confound standard interpretations. If it is a cultural norm not to use condoms, then deviating from this default may be more dread provoking when the prevalence of HIV is higher. In South Africa condom use is rare and HIV infection rates are high (MacPhaila and Campbell, 2001). It may be worth investigating whether anticipatory fear is a factor in condom avoidance.

Another issue that is plagued by inadequate preparation and widespread ignorance is planning for end-of-life care. This is a form of insurance that specifies the type of care one wants in the event of becoming terminally ill. Despite the benefits and professional support for the use of living wills in stating one's preferences for end-of-life care, less than 50% of adults have one (Carr and Khodyakov, 2007). If the default is to not prepare for end-of-life care then the model predicts that doing so would cause low anticipatory utility.

Bracha and Brown (2007) have a similar result in their model in which the decision maker simultaneously chooses beliefs and insurance. They show that an agent who buys more than full insurance will buy even more if the bad state worsens. An agent who buys less than full insurance will buy less if the bad state worsens. Our model makes a similar prediction. In the insurance example above, it is possible that increasing the probability of the bad state will decrease the amount an agent is willing to pay if he were planning on not buying insurance. Keeping the change in expected consumption utility fixed, our model makes the same prediction for decreasing the utility of the bad state, and increasing the probability of the bad state. These are different mechanisms in Bracha and Brown (2007). In their model there is a cost to distorting beliefs from the true probability, so updating about the probability of an event may incur costs, but updating about the utility conditional on the event is costless.

Mayraz (2009) has a model in which beliefs are distorted from the true probability by the payoffs. An "optimist" in his model will believe that the worse the state, the less likely the event. However, in his model, beliefs are monotonic in the true probability. Thus as the probability of the bad state increases, the agent will increase his belief and his insurance, although both will be biased downward.

The logic in this example and the information avoidance example is analogous. In terms of actions, deviating from the default is undesirable if future consumption is low. In terms of information, deviating from current beliefs is undesirable if future consumption is low. The result is generalized below. The anticipatory benefit to choosing a non-default action is decreasing in the expected consumption level.

Proposition 3 *Let $x \neq \bar{x}$. Keeping $w_i(x;b_1)-w_i(\bar{x};b_1)$ constant for all i, the difference $V_1(x;\bar{x},b_0,b_1)-V_1(\bar{x};\bar{x},b_0,b_1)$ is strictly increasing in $w_j(x;b_1)$.*

Proposition 3 says that, keeping the instrumental benefit of a deviating action fixed, the benefit to deviating is increasing in the expected consumption utility.[13] This implies that when expected consumption utility gets too low, the equilibrium action must be one in which attention is not shifted to the dimension. The intuition for this result is that people want to follow their default when even the best deviation produces a low expected consumption utility.

The "behavioral lock-in" result depends on knowing the level of expected consumption utility. Expected consumption utility in a dimension is not observable directly. If expected consumption utility in a dimension is strictly increasing in an observable variable, changes in this expected consumption may be observable. Moreover, as discussed in relation to Corollary 1 expected consumption utility may be indirectly observed via information preferences.

4.2 Information-Consumption Complementarity

The model predicts that information is a complement to consumption. When the DM receives information relevant to dimension i, he thinks about this dimension more, weights it more heavily in the utility function, and consequently trades off more consumption in other dimensions for consumption in dimension i. We can think of this as an "advertisement effect" in the sense of Becker and Murphy (1993) where advertisements are complements to consumption. Independent of the induced beliefs, a signal makes a person think more about the dimensions of consumption for which it is informative. Since the person is more intensely anticipating these dimensions of consumption, the person has a stronger preference for improving these dimensions.

An agent with $\gamma = \alpha = 0$ would also change his behavior upon obtaining information. However, what we show is that if the information-consumption complementarity is sufficiently strong, the DM may consume more in a dimension even after receiving information that suggests that the marginal benefit to consumption is lower than expected. An agent with $\gamma = 0$ would not behave in this manner. We illustrate with the following example.

Example : Consuming More When the Benefit Decreases

In this example we show that the DM increases consumption in a dimension after learning that the marginal benefit is low. Let $x \in \{0,1\}$ be future consumption in dimension i. There are two states. The

[13] If $w_j(x;b_1) \geq u_p$, then the difference is constant in $w_j(x;b_1)$. Once the level is above u_p, the DM wants to attend to the consumption and can freely do so without deviating from his plan.

consumption yields high consumption utility or low consumption utility $\Omega = \{h, l\}$. Initially, beliefs are that both states are equally likely. If the good is very enjoyable, then the benefit ω is h, otherwise the benefit is l, where $h > l$. There are two signals: revealing the state, and nothing, $\Sigma = \{r, n\}$. Choosing $\sigma = r$ reveals the true benefit of consuming the good. Both the state and the action only affect utility in dimension i. Expected consumption utility is $w_i(x; b_t) = E_{b_t}[\omega]x$.

First, we wish to know the relative benefit of consumption in the absence of information. Since there is no new information, the DM will choose $\bar{x} = x$. Utility will be $V_1(x; x, n, b_1) = \frac{h+l}{2}x$. The difference between consuming and not is $(h + l)/2$.

With information, $\sigma = r$, there are two possible defaults. Let us consider each in turn. If $\bar{x} = 1$, then $V_1(1; 1, r, b_1) - V_1(0; 1, r, b_1) = (1 + \gamma)\omega + \alpha u_p$. Doing the same for $\bar{x} = 0$, gives $V_1(1; 0, r, b_1) - V_1(0; 0, r, b_1) = (1 + \alpha + \gamma)\omega - \alpha u_p$.

In both of these, the DM's willingness to pay in utils is increasing in γ. If $\gamma > \frac{h+l+2\alpha u_p}{2l} - (1 + \alpha)$, then the DM's willingness to pay in utils is greater after learning that the marginal benefit is low than when receiving no information. The intuition is that the information itself increases the person's attention on dimension i. In this particular case, the attention-shifting effect of the signal overwhelms the actual content of the signal and so the person adjusts in the direction opposite to the realization of the signal.

The example illustrates the information-consumption complementarity. Loosely stated, information relevant to a dimension of consumption i increases demand for consumption in i. The complementarity works in both directions as indicated by Proposition 2 (when consumption increases in i, the willingness to pay for information about i increases). In order for behavior to actually exhibit increased consumption in i, there must be a tradeoff between consumption utility in i and another dimension.

One potential application of the information-consumption complementarity is that information can influence time preferences. If dimensions correspond to different time periods, then the complementarity will influence the weighting of different periods. Information about distant future episodes would cause the DM to behave as if he had a large discount factor. Information about the near future may make the DM appear to behave as if he were impatient.

The example illustrates that it is *possible* for the information-consumption complementarity to be observed. However, if the complementarity is weak, then it may be swamped out by the standard instrumental effect of information. In order to identify the information-consumption complementarity, we examine the behavior of the DM based on the timing of the receipt of information. To proceed, we define two different agents with identical preferences and beliefs, a prior-informed (PI) agent and a recently-informed (RI) agent. The PI agent begins with beliefs b_0^{PI} and receives a signal σ_n with no information, so his posterior $b_1 = b_0^{PI}$ for all realizations of the signal. The RI agent begins with beliefs b_0^{RI} and receives a signal σ and realization s such that his posterior is also $b_1 \neq b_0^{RI}$. Receiving the signal shifts attention to relevant dimensions. As a consequence, the two agents with identical preferences and identical beliefs, will behave differently. The two agents have attention shifted in different ways which will affect their anticipatory utility, and their behavior. The agent who became recently informed cares more about future consumption relevant to the information. To better identify the information-consumption complementarity, we compare a PI agent to an RI agent.

Proposition 4 *Let $x', x'' \in X$ such that $w_i(x'; b_t) > w_i(x''; b_t)$ and $w_j(x'; b_t) = w_j(x''; b_t)$ for all b_t and $j \neq i$. Furthermore let the signal σ only affect dimension i: $w_i(x; b_1) \neq w_i(x; b_0^{RI})$ for some $x \in X$, and $w_j(x; b_1) = w_j(x; b_0^{RI})$ for all $j \neq i$ and all $x \in X$. Then for a fixed $\bar{x}$, $V_1(x'; \bar{x}, b_1, b_0^{RI}) - V_1(x''; \bar{x}, b_1, b_0^{RI}) > V_1(x'; \bar{x}, b_1, b_0^{PI}) - V_1(x''; \bar{x}, b_1, b_0^{PI})$*

Proposition 4 says that an RI agent who received information that affected utility in dimension i will pay more to increase consumption in dimension i than the PI agent. The caveat is that the comparison is made when both agents have the same default. More generally, it may not be the case that these two agents choose the same default.

The timing of information here is critical. What is the difference between the PI and RI agents if both receive information before taking actions? We interpret the difference to be a matter of degree. A PI agent was informed long before the RI agent. This is admittedly a somewhat subjective matter. However, the model still does make a clear observable prediction. A person who was informed more recently will consume relatively more in the relevant dimensions than one who was informed less recently.

It is important to note that the information-consumption complementarity exists only when the receipt of information and the choice of consumption are contemporaneous. It is not the beliefs themselves that causes the complementarity, but the *change* in beliefs. Presumably, the PI agent at some point before period 1 received information and also had attention shifted to future consumption. However, by the time period 1 arrives with the consumption choice, attention had already returned to baseline and so the PI agent is unable to exploit the information-consumption complementarity.

An implication is that information campaigns will only have a temporary effect. The model predicts that there are two effects from an information campaign, the standard informational effect and the complementarity. Since the complementarity is temporary, the model predicts that the effect of a fixed-duration information campaign should decline over time. Murry, Jr. et al. (1993) study the effect of a paid advertising campaign to discourage drinking and driving targeted at young drivers. They found that the advertisement reduced fatal accidents for young drivers during the duration of the campaign and a few months later but the effect was only temporary.

5 Application: Advertising

In this section we apply the model to advertising. Advertisement as information has a long history in economics. See Bagwell (2007) for a broad survey. We model advertisement as certifiable information about the exogenous quality of an attribute of a multi-attribute good. In the language of Milgrom (1981), we consider a variant of the "Persuasion Game". According to the famous unraveling logic of Grossman (1981) and Milgrom (1981), the seller with highest quality will reveal the quality in order to separate from lower types and thus command a higher price. Consequently, any seller not revealing must be lower than highest quality. But then the seller with slightly lower quality will do strictly better by revealing. The logic continues until only the lowest type is left indifferent between revealing and not revealing. In the standard version of this game, the only sequential equilibrium is one in which all sellers with quality above the worst reveal their quality.

In our variant of this game, revealing the quality of an attribute will have two effects. First, it will have the standard effect to update the consumers' beliefs. If the quality is above average then disclosing the quality can increase the consumer's willingness to pay for the good. Second, through the information-consumption complementarity, the consumer will pay more attention to the attribute. More weight will be placed on the dimensions that have different expected consumption utility from expected consumption utility under the prior. As a result, the consumer will exhibit greater demand for attributes for which new information was provided. This effect is in the spirit of Becker and Murphy (1993) where advertisements are treated as complements with their respective good. Here, revealing the quality of an attribute of the good will increase the importance of that attribute to the consumer. The firm can use this complementarity to increase the price whenever the consumption utility in the dimension associated with the revealed attribute is greater than the consumption utility in the same dimension from not buying the good.

By the same logic, advertising attributes in which the consumption utility is greater from not buying can lower the price. The good in its entirety may be desirable but there may be undesirable attributes such as safety, maintenance, or inconveniences associated with using the good. We show that there exists a sequential equilibrium in which the firm will not reveal attributes and expected consumption utility in the dimension is substantially greater from not buying. In other words, the firm will shroud even "good news" about the attributes of their good that make the consumer worse off. A drug company will avoid advertising side effects of their drugs even if the side effects are better than average. The standard motivation to disclose high quality in this dimension is still present, but the information-consumption complementarity reduces the consumer's willingness to pay. The full disclosure equilibrium will also exist, but if the monopolist can choose which equilibrium to play, then the monopolist would prefer the shrouding equilibrium with any level of quality.

We then consider welfare in this framework. The firm's decision to shroud an attribute depends on whether the quality of the good for that attribute is sufficiently below the quality of not buying. However, welfare depends on whether the quality of the good is above the forgone utility of attention. These two comparisons are unrelated to each other and so in general, the firm will not be maximizing social welfare. There may be both too much or too little disclosure. In a richer framework there are two additional issues. First, consumers have access to outside information when there is too little advertisement. Second, if quality were endogenous, lack of disclosure may lead to an under-provision of quality.

When a second firm is added to the market we show that competition will increase the amount of disclosure. If the value of the goods are sufficiently high then there will be full disclosure in equilibrium. Each firm can increase its profit if it can increase the consumers' beliefs that its quality is above the competitor's. Since there is always a firm that has higher quality, there will always be one firm that has an incentive to disclose.

5.1 Setup

There is a monopolist and a single consumer. We begin with the timing followed by the monopolist's choices and preferences, and then the consumer's choices and preferences. The timing is illustrated in Figure 2. First, the consumer enters with beliefs b_0 about the distribution of quality and then decides his default action, then Nature chooses the quality of the good which is only observable to the monopolist, the monopoly chooses a reporting strategy and a price, the consumer buys the good or not, anticipatory utility and profit are realized, and finally quality is revealed and consumption utility is experienced in $t = 2$.

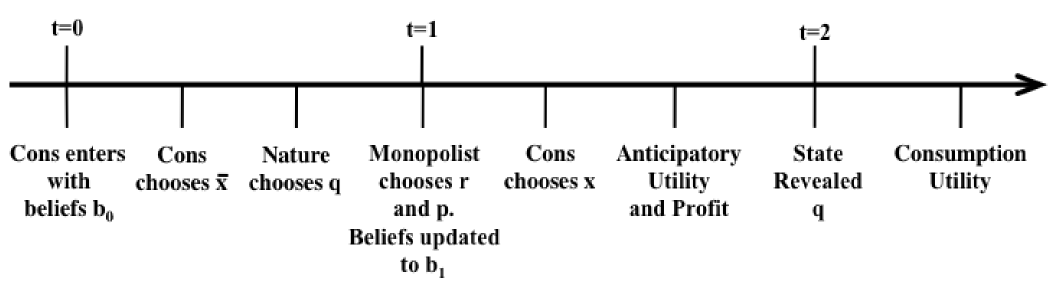

Figure 2: Game Timeline

The monopolist possess a single good which contains exogenous quality in I attributes. Each of these attributes corresponds to a dimension in the consumer's preferences. For example, if the good is heart medication, one attribute may be heart health, and another attribute might be the unintended side effect of bad digestion. An attribute in dimension i has quality $q_i \in \{q_i^1, ..., q_i^m\}$ ordered in increasing quality, and is known only to the monopolist. Each attribute may be either a good or a bad. We assume that not buying is equivalent to possessing a good with zero quality in every attribute, $q_i = 0\ \forall i$. Let q be the vector of quality. Quality in this game is the state, and all potential quality levels are the state space. Nature determines q from a commonly known distribution b_0. A critical assumption that we make is that q_i and q_j are independent for all dimensions i and j. As we explain later, independence is necessary for our results. The monopolist's strategy consists of two components, a report and a price. The monopolist may choose to certifiably report the exact quality of any number of attributes or not. Let the report in dimension i be to disclose or to shroud, $r_i \in \{d, s\}$. The full report conditional on quality is $r(q) \in \{\{d, s\}_{i=1}^I\}$. A price is any non-negative number $p(q) \in \mathbb{R}_+$. The monopolist has zero cost to producing the good, so the monopolist's profit is simply equal to the price, $\Pi = p$.

The consumer begins by choosing a default action, to buy or not buy, $\bar{x} \in \{B, N\}$ which we assume is observable by the monopolist.[14] The consumer has at most unit demand and would always purchase the good if $p = 0$ (for both defaults). The consumer has a degenerate choice of only one signal and the realization is determined by the monopolist's report. The consumer then chooses to buy or not buy $x \in \{B, N\}$ as a function of expected quality and price. We consider the consequence z to be possession of quality q. The action "buying" B given the state q produces the consequence, possession of quality q. The action "not buying" N given the state q produces the consequence, possession of quality 0. Technically, utility in dimension i takes the whole consequence vector q as an argument. Since utility in dimension i only depends on the quality of attribute i, we will switch notation slightly using the shorthand $u_i(q_i)$. Loosening the assumption from Section 3, we now allow both $u_i(q_i) \geq u_p$ and $u_i(q_i) < u_p$. Our welfare analysis will depend on these inequalities.

We assume that utility is quasi-linear in a numeraire good. If the price equals the consumer's expectation of the price, then less attention will be on the numeraire than if the price did not equal the consumer's expectation.[15] This attention-shifting to the numeraire via information about the price substantially complicates the analysis. To simplify, if the price of numeraire is $1+\nu$, where ν is a zero-mean random variable with arbitrarily small variance and revealed before the monopolist moves,

[14] We assume that the default is observable to make the analysis simple. If the monopolist does not observe the default, then we can think of the monopolist's price and report, and the consumer's choice of default as occurring simultaneously. This increases the number of equilibria but does not eliminate the equilibria to be described.

[15] It is worth mentioning that uncertainty over numeraire consumption is a function of the monopolist's choice. In all other parts of the model, both in this section and previous sections, uncertainty over consumption in a dimension is a function of the state and not choice variables.

then as long as ν never equals 0, γ attention will always shift to the numeraire dimension. This is admittedly an ancillary assumption about the environment but it simplifies the results and crisply helps to illustrate the intuition. Without this assumption, the basic result that the firm will shroud attributes remains true. Let $a_n(\bar{x}, x; b_0, b_1)$ be attention on the numeraire dimension, and let $-p$ be utility in the numeraire dimension. Total utility from buying and not buying given price p and report r are given by:

$$\begin{aligned} V_1(B; \bar{x}, b_0, b_1, p) = & -p + a_n(\bar{x}, B; b_0, b_1)(-p - u_p) \\ & + \sum_{i=1}^{I} E_{b_1}[u_i(q_i)] + a_i(\bar{x}, B; b_0, b_1)(E_{b_1}[u_i(q_i)] - u_p) \end{aligned} \tag{4}$$

$$V_1(N; \bar{x}, b_0, b_1, p) = -a_n(\bar{x}, N; b_0, b_1)u_p + \sum_{i=1}^{I} u_i(0) + a_i(\bar{x}, N; b_0, b_1)(u_i(0) - u_p) \tag{5}$$

5.2 Solution

The solution concept we use for the game is sequential equilibrium (Milgrom and Roberts, 1986; Okuno-Fujiwara et al., 1990; Osborne and Rubinstein, 1994). In this particular game, the sequential equilibrium will be equivalent to a perfect Bayesian equilibrium (Osborne and Rubinstein, 1994). A triple $((\bar{x}^*, x^*), (r^*, p^*), b_1(r^*))$ is a sequential equilibrium in pure strategies if it satisfies the following:

1. Monopolist maximizes profit: $\Pi((\bar{x}^*, x^*), (r^*, p^*), b_1(r^*)) \geq \Pi(((\bar{x}^*, x^*), (r, p), b_1(r)) \ \ \forall (r, p)$.
2. Consumer maximizes expected utility in both periods:

$$V_1(x^*; \bar{x}, b_0, b_1(r^*), p^*) \geq V_1(x; \bar{x}, b_0, b_1(r^*), p^*) \ \ \forall x \in X,$$

$$V_0(\bar{x}^*, \sigma; x^*, b_0, p^*) \geq V_0(\bar{x}, \sigma; x^*, b_0, p^*) \ \ \forall \bar{x} \in X.$$

3. The consumer has rational beliefs: If $r_i = s$, $\Pr_{b_1}(q_i|r) = \frac{\Pr_{b_0}(r_i = s|q_i)\Pr_{b_0}(q_i)}{\sum_j \Pr_{b_0}(r_i = s|q_j)\Pr_{b_0}(q_j)}$.
4. The consumer has consistent beliefs: If $r_i = d$ and realized quality in i is q_i', then $\Pr_{b_1}(q_i'|r) = 1$.

We begin by solving backwards. The consumer will buy if and only if $V_1(B; \bar{x}, b_0, b_1, p) \geq V_1(N; \bar{x}, b_0, b_1, p)$. The monopolist, knowing this will set prices to extract all expected surplus from the purchase,

$$\begin{aligned} p^*(q) = & \frac{1}{1 + a_n(\bar{x}, B; b_0, b_1)} \Bigg(\big(a_n(\bar{x}, N; b_0, b_1) - a_n(\bar{x}, B; b_0, b_1)\big) u_p \\ & + \sum_{i=1}^{I} E_{b_1}[u_i(q_i)] + a_i(\bar{x}, B; b_0, b_1) \left(E_{b_1}[u_i(q_i)] - u_p\right) - u_i(0) - a_i(\bar{x}, N; b_0, b_1)(u_i(0) - u_p) \Bigg). \end{aligned} \tag{6}$$

The multiplier on the left is the inverse of how much the consumer weights the numeraire consumption. The first term in the parentheses represents the relative foregone attention from attending to numeraire consumption. The first two terms in the summation are the utility from purchasing the good, and the second two terms are the utility from not purchasing the good. This leaves the consumer with the utility from not buying, $V_1(N;\bar{x},b_0,b_1,p)$. For this reason, the consumer's utility is always weakly greater if the consumer plans to not buy, $\bar{x}=N$. To remind the reader, attention will not shift to low level attributes when $\bar{x}=x$ and attention always shifts to high level attributes independent of $\bar{x}$. Therefore, if the consumer comes to the marketplace expecting to get $V_1(N;\bar{x},b_0,b_1,p)$, then choosing the habit of not buying $\bar{x}=N$ is optimal. The interpretation is that relative to forming the habit of buying, this increases the consumer's reservation utility. "Non-habitual" consumers have a higher outside option so they command better deals.[16]

Remaining to be solved is the monopolist's report and the consumer's beliefs about quality. With standard preferences the only sequential equilibrium is full disclosure (Grossman, 1981; Milgrom, 1981; Okuno-Fujiwara et al., 1990). We now construct a sequential equilibrium in which the firm shrouds some attributes from the consumer. We begin by considering the consumer's side of the problem. Suppose the monopolist will shroud attribute i, $r_i=s$ for any realization of q_i. If this is the case then the consumer can infer nothing about the quality from observing $r_i=s$, because all types are shrouding, and so $E_{b_1}[q_i]=E_{b_0}[q_i]$. Since expected quality does not change in the dimension, no attention will shift to the dimension. The monopolist's incentive to deviate is based on its quality. If the monopolist has the highest quality, $q_i=q_i^m$, then the incentive to separate from the lower types is greatest. The benefit from deviating for this monopolist is

$$p^*_{r_i=d}-p^*_{r_i=s}=\frac{1}{1+\alpha+\gamma}\big((1+\alpha)\left(u_i(q_i^m)-E_{b_0}[u_i(q_i)]\right)+\gamma\left(u_i(q_i^m)-u_i(0)\right)\big) \qquad (7)$$

If the consumption utility from not buying in dimension i is substantially higher than buying $\gamma(u_i(0)-u_i(q_i^m))>(1+\alpha)(u_i(q_i^m)-E_{b_0}[u_i(q_i)])$ then the firm will shroud, even with the highest quality.

Proposition 5 *There exists a sequential equilibrium in which the monopolist shrouds in dimensions j if and only if $\gamma(u_j(0)-u_j(q_j^m))\geq(1+\alpha)\big(u_j(q_j^m)-E_{b_0}[u_j(q_j)]\big)$.*

The left hand side of the expression is the utility difference between buying and not buying in dimension i weighted by the shift of attention from information, and it represents the loss to the firm from advertising a low attribute. The right hand side of the expression is the utility difference from

[16] If $\bar{x}$ is unobservable to the monopolist, there will be another set of pure strategy equilibria in which the consumer chooses the default $\bar{x}=B$, the monopolist will set prices to leave the consumer indifferent between buying and not, and the consumer will buy. In this set of equilibria, the monopolist will obtain a greater fraction of the surplus. More pertinently, the monopolist's choice to shroud or reveal will remain unaffected.

the best quality and the unconditional average quality weighted by the shift of attention from actions, and it represents the benefit to the highest-quality monopolist from separating from types who have low quality. Notice that if there is no information-consumption complementarity, that is $\gamma = 0$, the inequality is never satisfied, and so the only equilibrium is full disclosure for all dimensions.

The intuition is that by disclosing, the monopolist shifts attention to the attribute. This can be effective at increasing the price when the consumption utility from buying is much higher than not buying. However, the firm does not wish to draw attention when the attribute is a drawback to consumption. For example, a drug company may advertise the primary effect of their drug, an attribute for which quality is high relative to no drug, but the firm may not wish to advertise the side effect of the drug, an attribute for which quality is worse than no drug. If the quality in several dimensions are correlated, then disclosing quality in one dimension will change expected quality in the other dimensions, and thereby shift attention to these dimensions.

Equilibria in which there is full disclosure in dimension i also exist. To see this suppose the consumer believes the monopolist will disclose attribute i for all q_i and that any shrouding comes from types possessing the lowest quality q_i^1. Under these beliefs, shrouding will still change beliefs about quality from the prior beliefs, and so $w_i(B; b_1) \neq w_i(B; b_0)$ and thus attention will shift to dimension i. Since attention shifts to dimension i for either disclosing or shrouding, the monopolist has no incentive to shroud.[17] In other words there is an "attentional externality". As long as at least one type discloses, attention will shift for all types, even if they do not disclose. Shrouding will cause the consumer to infer quality is the lowest $q_i = q_i^1$. The standard unraveling logic applies. Whenever there exist two sets of equilibria, one in which q_i is shrouded and one in which q_i is disclosed, the monopolist will prefer the shrouding equilibria for all realized q_i. The condition for preference is the same condition that allows shrouding in equilibrium, equation (7).

While our model predicts one reason that the unraveling argument may fail, the existing literature has uncovered several others (Okuno-Fujiwara et al., 1990). We consider these and contrast them with the predictions of our model. There are several conditions in which full disclosure will break down into partial disclosure, in which only types with quality above a threshold will disclose. First off, the unraveling argument fails if disclosure is costly. If costs are sufficiently high there will be partial disclosure or no disclosure. The tradeoff for the firm is between the benefit of separating and the cost of disclosure. All types below a threshold will have little benefit to separating and thus prefer to pool. In the psychological model of Eyster and Rabin (2005) agents under-appreciate the correlation of other agents' behaviors with their types. It is as if the observer mistakenly believes

[17] This prediction depends on the assumption that the attention function is discrete. With a continuous attention function we speculate that the full disclosure equilibrium would become a partial unraveling equilibrium.

that some types who disclose shroud, and some who shroud disclose. The observer infers less from shrouding than a rational agent would infer, an thus the consumer's expected utility after observing pooling is mistakenly higher than the actual expected utility from pooling. Types with low quality will exploit this bias by pooling. Hirshleifer et al. (2002) present a psychologically motivated model in which observers behave in two non-standard ways which the authors collectively label as "limited attention". There is some probability that observers ignore the disclosure of the sender, and there is some probability that observers fail to update their beliefs conditional on no disclosure. Under these assumptions the observer will over-estimate quality conditional on no disclosure, and thus there will only be partial disclosure since senders with low quality can do better by being silent. A breakdown of common knowledge about quality could prevent unraveling too. For example, if the monopolist only knew the quality of the product with some probability, then silence in equilibrium may imply that the firm does not know the quality. The consumer's expected utility after observing pooling will be higher than the expected utility from those who *choose* to pool, because there will be some high types who are forced to pool since they are unable to signal. As a result, types with low quality will choose to pool.

We can distinguish from all four of these possibilities since our model will not allow for partial disclosure. The assumption that attention shifts discretely drives our result. If the shift in attention were increasing in the change of expected consumption utility, then partial disclosure may occur in equilibrium. But more importantly, our model predicts that the monopolist will simultaneously disclose some attributes and shroud others. These other theories cannot make this prediction unless there are different costs for revealing different attributes, different degrees of mis-prediction for different attributes, or different degrees of knowledge for different attributes, respectively. Furthermore, there are some situations where common knowledge would seem appropriate (e.g. drug companies are required by law to conduct clinical trials for the FDA before they can go to market).[18]

Finally, the result can breakdown if utility is not monotonic in quality, or tastes in the population are heterogenous. These are reasonable concerns but we do not think it is too much of a stretch to assume that there are some attributes for which all individuals would prefer higher quality. It is hard to imagine a consumer who would prefer a drug to cause more nausea.

5.3 Welfare

We now analyze welfare effects of the advertisement. Care should be made in interpreting these results. There are many facets to advertising and we are looking at one, the effect of advertisements on directing

[18]Information available at http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/default.htm

attention to future consumption. Advertisements may have additional benefit such as providing useful information and they may have additional costs such as stealing consumers away from competitors (Bagwell, 2007). In the following analysis we ignore all other potential concerns. We take the social welfare function to be the sum of consumer utility and profit converted from money to utils. Since one numeraire equals $1+\alpha+\gamma$ utils, $SW = V_1 + (1+\alpha+\gamma)\Pi$. In this subsection, we will compare welfare under mandatory disclosure and prohibited advertisement to welfare under equilibrium.

We begin by comparing utility in an equilibrium in which there is full shrouding in dimension i, to a regime in which the monopolist is required to disclose. The reporting strategy is forced to have $r_i(q) = d$ for all q. First we analyze the effect on the consumer. Disclosing shifts attention to dimension i. In equilibrium, the consumer buys but the price is such that he is left indifferent between buying and not buying and so gets the utility $V_1(N; N, b_0, b_1, p^*)$. The effect of learning about quality on the consumer is $\Delta V_1 = \gamma(u_i(0) - u_p)$. This can be either negative or positive. The effect of disclosing quality for the firm will be the gain from separating plus the gain or loss from the information-consumption complementarity $\Delta\Pi = \frac{1}{1+\alpha+\gamma}(1+\alpha)(u_i(q_i) - E_{b_0}[u_i(q_i)]) + \gamma(u_i(q_i) - u_i(0))$. Thus the change in social welfare will be $\Delta SW = \Delta V_1 + (1+\alpha+\gamma)\Delta\Pi = (1+\alpha)(u_i(q_i) - E_{b_0}[u_i(q_i)]) + \gamma(u_i(q_i) - u_p)$. If we take the average over all quality levels we get $E_{b_0}[\Delta SW] = \gamma(E_{b_0}[u_i(q_i)] - u_p)$. To perform the reverse comparison, that is comparing utility in an equilibrium in which there is full disclosure in dimension i, to a regime in which the monopolist is prohibited from disclosing quality in dimension i, we simply reverse the sign of ΔSW.

So whenever the attribute's average expected consumption utility is greater than the forgone utility, full disclosure of the dimension on average produces higher welfare than full shrouding. Intuitively, shifting the consumer's attention to an attribute that is pleasurable to anticipate increases welfare. For example, receiving an advertisement about one's hobby may be pleasurable, because even if the quality of the product is low, one thinks about an enjoyable topic. When expected consumption utility is less than the forgone utility, but the firm finds it profitable to disclose, welfare is decreased. The firm draws attention to attributes that may be better than not buying the good, but that are unpleasant to think about. For example, a monopolist advertising the quality of heart medication will draw attention to a consumer's debilitating health, something which the consumer may not enjoy thinking about. In the language of Becker-Murphy, this advertisement is a bad. Even though the advertisement increases demand for the drug, the consumer would prefer not to think about such matters.

There is little reason to expect that there is a systematic relationship between the difference of the utilities of buying and not buying $(u_i(q_i) - u_i(0))$, and whether the dimension is pleasant to anticipate $(u_i(q_i) - u_p > 0)$. So the firm's advertising decisions are essentially orthogonal to consumer welfare. It is then difficult to say whether there is too much or too little advertisement. When there is too much advertisement, consumers are being “forced” to attend to dimensions that they would rather not think about. A typical consumer is barraged with advertisements that he did not choose: from billboards, to advertisements on buses. If a consumer with a heart condition receives information about a heart drug, this draws the consumer’s attention to his health which may be depressing. When there is too little advertisement, consumers want to think about positive things but do not have any information to shift attention. For example, a person living in Hawaii wants to receive more information about beaches and tropical wilderness, but a travel agency based in Hawaii cannot profit from this so they instead advertise the skyline of New York City.

In a natural variation of the model in which consumers have cheap sources of information other than advertisers, a resident of Hawaii could freely acquire information about beaches. The consumer could get information from alternative non-adverisement sources. If information that can shift attention to any dimension is freely available to the consumer, then it must be the case that there is too much advertisement. The consumer obtains information for all dimensions in which $E_{b_0}[u_i(q_i)] \geq u_p$. Only the undesirable advertisements for which $E_{b_0}[u_i(q_i)] < u_p$ will have an additional effect.

However, in this simple world of exogenous quality and zero costs the only role of disclosure is for the firm to command a higher price. Efficiency requires that the consumer buys, and that the firm advertises an attribute if and only if $u_i(q_i) \geq u_p$. However, when quality is endogenous and costly to produce there may be efficiency lost from too little advertisement. Suppose nature exogenously chooses the cost function (instead of quality) that the monopolist uses to produce quality. For dimensions in which the utility from buying is substantially less than the utility from not buying, not only will the monopolist shroud, but because the monopolist shrouds there will be zero investment in quality for that dimension. A monopolist that invests a positive amount may reveal or not, but revealing will make the price lower. So the monopolist has no profitable way to reveal quality and hence no incentive to produce quality. This can lead to a market inefficiency. The consumer benefit to higher quality may exceed the costs but does not want his attention drawn to it. A consumer may want the drug company to reduce the side effects, but does not want to have to “think about it”. We can call this an “attentional transaction cost”. The monopolist will be unwilling to communicate information that draws attention to displeasurable attributes (unless the expected consumption utility difference between investing optimally and zero investment is large, in which case the firm will choose to invest optimally and disclose). This suggests a role for third parties to monitor and ensure quality within an industry without directly informing the consumer about quality. The consumer will infer that quality is higher by the presence of the monitor but need not learn the details. The firm would benefit as well because quality and hence higher prices would be higher.

There is empirical evidence that mandatory disclosure increases quality. Jin and Leslie (2003) find that disclosure of restaurant hygiene reports in Los Angeles increases hygiene. In 1997 Los Angeles passed an ordinance that required all restaurants to prominently display A, B, or C hygiene grade cards in their window. After implementation, inspection scores increased by 5.3%, grade A restaurants made 5.7% more revenue (average revenue increased as well), and there was a 20% decrease in food-borne illness related hospitalizations. Additionally, Bennear and Olmstead (2008) found that mandatory disclosure of drinking water violations in Massachusetts reduced total violations by between 30–44% and 40–57% for severe violations. Our analysis suggests that, while quality investment is good for both consumer and firm, disclosure may also hurt both because consumers do not want to think about restaurant hygiene violations when they are eating their dinner. Hypothetically, verification by a third party may both insure high quality without concerning consumers.

5.4 Extension: Duopoly

In this extension, we consider the role of competition on unraveling. As in Milgrom and Roberts (1986), one may suspect that competition between informed parties will elicit all information. To address this issue, we use a Hotelling model in which two firms compete for a continuum of consumers with transportation costs. We show that competition can completely eliminate shrouding. Firms obtain more customers by revealing attributes for which their product is better than their opponent's product. If the goods are valuable enough so that all consumers buy, there can only be full disclosure in equilibrium.

We extend the basic setup in two major ways. There is now a continuum of consumers, each of whom lives at a point $\rho \in [0, 1]$ on a linear city with uniform density. As previously, consumers have at most unit demand. There are now two firms, firm 0 and firm 1 located at each end of the city. The consumers must pay a transportation cost to purchase the good. One can also interpret ρ as a taste parameter. The transportation cost for purchasing from the first firm for a consumer at ρ is ρt and the cost for purchasing from the second firm is $(1 - \rho)t$. We assume these costs are dimensionless with no anticipatory utility. Nature chooses quality for each firm, and quality of the firms q and k respectively, are independent. The firms have common knowledge about each other's quality, but the consumer only knows the distribution from which quality is drawn. The timing is the same, and the firms move simultaneously. Denote $p^0(q)$ and $r_i^0(q)$ to be firm 0's price and reporting strategy for attribute i, and likewise $p^1(k)$ and $r_i^1(k)$ for firm 1. Denote action B^0 and B^1 to indicate that the consumer purchases the good from firm 0 or 1 respectively.

Preferences and profit are slightly different now. To be explicit, a consumer's utility is

$$\begin{aligned}
V_1(B^0;\bar{x},b_0,b_1,p,\rho) = &-\rho t - p^0 + a_n(\bar{x},B^0;b_0,b_1)(-p^0-u_p) \\
&+\sum_{i=1}^{I} E_{b_1}[u_i(q_i)] + a_i(\bar{x},B^0;b_0,b_1)(E_{b_1}[u_i(q_i)]-u_p)
\end{aligned} \tag{8}$$

$$\begin{aligned}
V_1(B^1;\bar{x},b_0,b_1,p,\rho) = &-(1-\rho)t - p^1 + a_n(\bar{x},B^1;b_0,b_1)(-p^1-u_p) \\
&+\sum_{i=1}^{I} E_{b_1}[u_i(k_i)] + a_i(\bar{x},B^1;b_0,b_1)(E_{b_1}[u_i(k_i)]-u_p)
\end{aligned} \tag{9}$$

$$\begin{aligned}
V_1(N;\bar{x},b_0,b_1,p,\rho) = &-a_n(\bar{x},N;b_0,b_1)u_p \\
&+\sum_{i=1}^{I} u_i(0) + a_i(\bar{x},N;b_0,b_1)(u_i(0)-u_p)
\end{aligned} \tag{10}$$

The firms' profits are $\Pi^0 = p^0\rho^0$ and $\Pi^1 = p^1(1-\rho^1)$ where ρ^0 is the rightmost consumer that purchases from firm 0 and ρ^1 is the leftmost firm that purchases from firm 1.

To solve, when all consumers buy, then $\rho^0 = \rho^1$. Let us call this point $\rho^* \equiv \rho^0 = \rho^1$. This means a consumer with ρ^* is indifferent between the two goods $V_1(B^0;\bar{x},b_0,b_1,p,\rho^*) = V_1(B^1;\bar{x},b_0,b_1,p,\rho^*)$. We will *assume* the default is $\bar{x} = N$. If we do not, there will be a discontinuity in the consumers' willingness to pay at $\rho = 1/2$.[19] This assumption may not be consistent with the setup in which the default is exogenous, but it is plausible to assume that the good in question is not a routine or habitual purchase.

First we find $\rho^* = \frac{1}{2t}(1+\alpha+\gamma)(p^1-p^0) + \frac{1}{2} + Y$, where

$$Y \equiv \sum_{i=1}^{I} \frac{1}{2t}\Big(\big(1+a_i(N,B^0;b_0,b_1)\big)E_{b_1}[u_i(q_i)] - \big(1+a_i(N,B^1;b_0,b_1)\big)E_{b_1}[u_i(k_i)]\Big).$$

The next step is to solve for the profit-maximizing price for each firm. Taking the first-order condition for firm 0 with respect to price yields the best-response function for price $p^{0*} = \frac{t}{1+\alpha+\gamma}(\frac{1}{2}+Y) + \frac{1}{2}p^1$. Similarly, $p^{1*} = \frac{t}{1+\alpha+\gamma}(\frac{1}{2}-Y) + \frac{1}{2}p^0$. Solving gives $p^{0*} = \frac{t}{1+\alpha+\gamma}(1+\frac{2}{3}Y)$ and $p^{1*} = \frac{t}{1+\alpha+\gamma}(1-\frac{2}{3}Y)$.

The point of this exercise is to see how duopoly affects disclosure. The key term in these expressions is Y which is a weighted quality difference between firm 0 and firm 1's goods. Notice that a firm that has higher quality in dimension i than its competitor can only increase its price by disclosing. By disclosing, more attention shifts to this attribute via the information-consumption complementarity. This magnifies the effect of the difference in quality on the price. The higher quality firm imposes

[19] The default of $\bar{x} = N$ will not be optimal. A consumer would prefer to set the default to buy from the firm from which they expect to buy. But a consumer does not know which firm they will buy from at the time at which they choose the default. If consumers choose the optimal default, all $\rho \le 1/2$, will choose $\bar{x} = B^0$ and and the rest choose $\bar{x} = B^1$. This generates the discontinuity.

an "attentional externality" on the lower quality firm. The weakly lower quality firm cannot avoid attention being drawn to its attribute even if it shrouds. The firm with weakly lower quality then has no attention-based incentive to shroud. The standard unraveling logic will apply. A firm with weakly lower quality equal to the competitor would disclose, so a firm with weakly lower quality that is slightly lower than the competitor would disclose, and so on.

Proposition 6 *There exists a constant* $\chi > 0$ *such that if* $\sum_{i=1}^{I} u_i(q_i) - u_i(0) \geq \chi$ *and* $\sum_{i=1}^{I} u_i(k_i) - u_i(0) \geq \chi$*, then* $r_i^0(q_i) = d$ *and* $r_i^1(k_i) = d$ *for all* i*, all* $q_i > q_i^1$*, and all* $k_i > k_i^1$ *in any equilibrium.*

Proposition 6 says that if the consumption value of the goods are sufficiently high, then any equilibrium will result in full disclosure. This is a possibility result, meaning that it is possible for the equilibria to have only full disclosure although this need not be the case. The critical assumption that leads to unraveling is that the good is sufficiently valuable such that the two best choices for all consumers are B^0 and B^1, as opposed to buying nothing at all. When the quality of the two goods is high enough so that all consumers are assured to buy, the firm with higher quality has an incentive to disclose and the firm with lower quality must disclose in order to not be thought lowest quality. The key difference from the monopoly case is that there is always a firm that can profit by disclosing since one firm always has higher quality than the other.

If there are some consumers in the middle of the city that do not buy, then the model reduces to two local monopolies. Disclosing attributes for which the utility from not buying is sufficiently larger than buying increases the appeal of not buying, and thus can induce some consumers near the center to not buy. In other words, it is possible that if some attributes are shrouded, all consumers buy, but if one of those shrouded attributes unravels, some consumers will not buy. Under this circumstance, shrouding may be preferred by both firms.

Even though it is possible to have only full disclosure equilibria, it is not guaranteed by the presence of competition alone. Casually speaking, the competitive effect of winning over consumers from the other firm increases disclosure, but the need to keep the consumer interested in buying the product at all may limit disclosure.

This prediction contrasts with the prediction of Hotz and Xiao (2006). They present a duopoly Hotelling model with quality disclosure in which the consumer's preferences are monotonic but heterogeneous for quality. When preferences for quality and location on the line are correlated in such a way that those who have a greater demand for quality are located near the lower quality firm, disclosure may lead to a price competition that lowers profits for both firms. As a result, they make the opposite prediction that we make: they predict there will be less disclosure under duopoly.

6 Conclusion

Our goal in this paper has been to introduce a formal model of attention and anticipatory utility that can resolve puzzling human behavior. In Sections 4 and 5 we showed that the model produces plausible behavior that differs from standard theory in economically important ways. We conclude by discussing possible extensions to the model.

When there are two rounds of anticipatory utility, the decision maker will exhibit interesting time-inconsistent behaviors. First, information can be used as a commitment device. The first period self can use information to incentivize the second period self through the information-consumption complementarity. Second, the interaction between attention and anticipation also leads to motivated belief (Carrillo and Mariotti, 2000; Bénabou and Tirole, 2002). This manifests in two ways. First, when the expected consumption utility of a dimension is below a threshold, there is a demotivation effect: the decision maker will avoid information, and defer to the default in order to not think about a depressing dimension. To avoid this demotivation effect, the decision maker will prefer information that keeps expected consumption above the threshold. This results in subtle information preferences where the decision maker cares about the type and timing of information. For example, if the default for a dissertation-writing graduate student is to spend time teaching instead of research, the student may avoid information about the job market in order to avoid possible future low morale and low research productivity following a low signal. The student will also work hard to keep utility above the threshold.

Modeling mutually exclusive events as different consumption dimensions may be a fruitful extension to risk preferences. Receiving information about the likelihood of surviving a plane crash while standing in the airport security line makes one feel uneasy. One can think of this as attention shifting to the event "crashing". Information about low-probabilty events would cause them to be overweighted. As a consequence, the DM may buy lottery tickets and still purchase excessive insurance.

Furthermore, it is plausible, in a dynamic setting, that the DM may systematically mispredict his future utility with a bias towards his current attention-shifted preferences. This would combine projection bias (Loewenstein et al., 2003) with our model. A DM with projection bias would incorrectly believe that the information-consumption complementarity is permanent. This would produce an effect similar to the focusing illusion (Kahneman et al., 2006) where "Nothing in life is quite as important as you think it is while you are thinking about it."

Appendix: Proofs

Proof of Proposition 1

Define $x^c \equiv arg\, max_{x\in X} \sum_{i=1}^{I} w_i(x; b_1)$. Remember, $w_i(x; b_1) < u_p$ for all i, x. First we show $w_i(\bar{x}^*(b_0); b_1) = w_i(x^*; b_1)$. We backwards induct to the decision node for choosing $\bar{x}$. When $\bar{x} = x^*$, then $V_0(x^*, \sigma; x, b_0) = \sum_{i=1}^{I} w_i(x^*; b_0)$. When $\bar{x} \neq x^*$, then for dimensions $j \in J$ that have $w_j(\bar{x}) \neq w_j(x^*)$ and dimensions $k \in K$ that have $w_k(\bar{x}) = w_k(x^*)$,

$$V_0(\bar{x}, \sigma; x, b_0) = \sum_{j\in J}(1+\alpha)w_j(x^*; b_0) - \alpha u_p + \sum_{k\in K} w_k(x^*; b_0).$$

Since $w_j(x^*; b_0) < u_p$, choosing $\bar{x}^*(b_0)$ such that $w_i(\bar{x}^*(b_0); b_0) = w_i(x^*(b_0, b_0); b_0)$ for all i, is preferred. No attention shifts and so $V_0(\bar{x}^*(b_0), \sigma; x, b_0) = \sum_{i=1}^{I} w_i(x^*; b_0)$ and by definition the maximizer is x^c.

Proof of Proposition 2

Let $\bar{x_1}^*$ and x_1^* denote optimal default and actions conditional on receiving the information σ and let $\bar{x_2}^*$ and x_2^* be optimal default and actions conditional on receiving no information σ_n.

$$\begin{aligned} V_0(\bar{x_1}^*, \sigma; x_1^*, b_0) - & V_0(\bar{x_2}^*, \sigma_n; x_2^*, b_0) = \\ & E_{b_0}\left[\sum_{i=1}^{I}(1 + a_i(\bar{x_1}^*, x_1^*; b_0, b_1))w_i(x_1^*; b_1) - a_i(\bar{x_1}^*, x_1^*; b_0, b_1)u_p\right] \\ & \sum_{i=1}^{I} -(1 + a_i(\bar{x_2}^*, x_2^*; b_0, b_0))w_i(x_2^*; b_0) + a_i(\bar{x_2}^*, x_2^*; b_0, b_0)u_p. \end{aligned}$$

We rewrite this expression putting all the constant terms in K.

$$=E_{b_0}\left[(1 + a_i(\bar{x_1}^*, x_1^*; b_1, b_0))w_i(x_1^*; b_1)\right] - (1 + a_i(\bar{x_2}^*; x_2^*; b_0, b_0))w_i(x_2^*; b_0) + K$$

Since $w_i(x_2^*; b_0) < u_p$, $a_i(\bar{x_1}^*, x_1^*, b_1, b_0) = \gamma$ or $a_i(\bar{x_1}^*, x_1^*, b_1, b_0) = \alpha + \gamma$ and $a_i(\bar{x_2}^*, x_2^*, b_0, b_0) = 0$. In this case the difference is increasing $w_i(x_1^*; b_1)$.

Proof of Corollary 1

$$
\begin{aligned}
V_0(\bar{x},\sigma;x,b_0)-&V_0(\bar{x},\sigma_n;x,b_0)\\
&= E_{b_0}\left[\sum_{i=1}^{I}(1+\gamma)w_i(x;b_1)-\gamma u_p\right]-\sum_{i=1}^{I}w_i(x;b_0)\\
&= \sum_{i=1}^{I}\gamma(w_i(x;b_0)-u_p)
\end{aligned}
$$

This expression is strictly increasing in $w_j(x;b_0)]$.

Proof of Proposition 3

Utility from x and $\bar{x}$ are respectively $V_1(x;\bar{x},\sigma,b_1)=\sum_{i=1}^{I}(1+a_i(\bar{x},x;b_0,b_1))w_i(x;b_1)-a_i(\bar{x},x;b_0,b_1)u_p$ and $V_1(\bar{x};\bar{x},\sigma,b_1)=\sum_{i=1}^{I}(1+a_i(\bar{x},\bar{x};b_0,b_1))w_i(\bar{x};b_1)-a_i(\bar{x},\bar{x};b_0,b_1)u_p$. The difference is

$$
(1+a_j(\bar{x},x;b_0,b_1))w_j(x;b_1)-(1+a_j(\bar{x},\bar{x};b_0,b_1))w_j(\bar{x};b_1)+K
$$

where K is a constant (for the purposes of varying w_j).

Since $w_j(\bar{x};b_1)<u_p$, the difference can be rewritten as $\alpha w_j(x;b_1)+(1+a_j(\bar{x},\bar{x};b_0,b_1))[w_j(x;b_1)-w_j(\bar{x};b_1)]+K$ because $a_j(\bar{x},x;b_0,b_1)=\alpha+a_j(\bar{x},\bar{x};b_0,b_1)$. One can see that the difference is increasing in $w_j(x;b_1)$ while keeping $w_j(x;b_1)-w_j(\bar{x};b_1)$ constant.

Proof of Proposition 4

$$
\begin{aligned}
V_1(x';\bar{x},b_1,b_0^{RI})-V_1(x'';\bar{x},b_1,b_0^{RI})&-V_1(x';\bar{x},b_1,b_0^{PI})+V_1(x'';\bar{x},b_1,b_0^{PI})\\
=&\gamma(w_i(x';b_1)-u_p)-\gamma(w_i(x'';b_1)-u_p)\\
=&\gamma(w_i(x';b_1)-w_i(x'';b_1))>0
\end{aligned}
$$

Proof of Proposition 5

Assume $u_j(0)-u_j(q_j^l)\geq\frac{1+\alpha}{\gamma}(u_j(q_j^l)-E_{b_0}[u_j(q_j)])$ for at least one dimension $j\in J$. We construct a strategy profile and show it is a sequential equilibrium. The monopolist chooses $r_i(q_i)=d$ for all $i\notin J$

and all q_i, and $r_j(q_j) = s$ for all $j \in J$ and all q_j. The price as given in the text

$$p^*(q) = \frac{1}{1 + a_n(\bar{x}, B; b_0, b_1)} \Big(\big(a_n(\bar{x}, N; b_0, b_1) - a_n(\bar{x}, B; b_0, b_1)\big) u_p + \sum_{i=1}^{I} E_{b_1}[u_i(q_i)] + a_i(\bar{x}, B; b_0, b_1) \left(E_{b_1}[u_i(q_i)] - u_p\right) - u_i(0) - a_i(\bar{x}, N; b_0, b_1)(u_i(0) - u_p) \Big).$$

The consumer chooses $\bar{x} = N$, and at his second information set chooses $x = B$ if and only if $V_1(B; N, b_0, b_1, p) \geq V_1(N; N, b_0, b_1, p)$, otherwise $x = N$. The consumer's equilibrium beliefs are $\Pr_{b_1}(q_j | r_j = s) = \Pr_{b_0}(q_j)$ for dimensions $j \in J$, and if the realized value is q_i' then $\Pr_{b_1}(q_i' | r_i = d) = 1$ and $\Pr_{b_1}(q_i | r_i = d) = 0$ for all other $q_i \neq q_i'$. The off-the-equilibrium-path beliefs for dimensions are $j \in J$, $\Pr_{b_1}(q_j' | r_j = d) = 1$ if the realized value is q_j' and $\Pr_{b_1}(q_j | r_i = d) = 0$ for all other $q_j \neq q_j'$. The consumer has skeptical beliefs for dimensions $i \notin J$, believing the quality to be the lowest possible after observing shrouding, $\Pr_{b_1}(q_i^1 | r_i = s) = 1$, and $\Pr_{b_1}(q_i | r_i = s) = 0$ for all $q_i \neq q_i^1$.

Now we check the four conditions for sequential equilibrium. This price maximizes the firm's profit. Notice charging a higher price would prevent sale of the good. Disclosing any attribute $j \in J$ would lower the price, by construction. Shrouding any attribute $i \notin J$ would also lower the price because the consumer would believe it to be lowest quality. So the firm is profit maximizing. The consumer is indifferent between buying and not buying. If the consumer switched the default to buying, equilibrium utility would go down, so that would be suboptimal. So the consumer is maximizing. The consumer's equilibrium beliefs and out-of-equilibrium beliefs are rational and consistent by construction.

Proof of Proposition 6

First we show existence of a full disclosure equilibrium. Then we show that for all such equilibria shrouding is not possible when the value of the goods are sufficiently large. When the value of the good is sufficiently large so that all consumers buy, the equilibrium prices as given in the text are $p^{0*} = \frac{t}{1+\alpha+\gamma}(1 + \frac{2}{3}Y)$ and $p^{1*} = \frac{t}{1+\alpha+\gamma}(1 - \frac{2}{3}Y)$. Since we are constructing a full disclosure equilibrium $r_i^0 = r_i^1 = d$ for all i, all quality levels. A consumer will choose B^0 if and only if $V_1(B^0; N, b_0, b_1, p^0, p^1, \rho) > V_1(B^1; N, b_0, b_1, p^0, p^1, \rho)$, otherwise B^1. The consumer believes the quality is the realized quality upon disclosure, and believes quality is the lowest if shrouded. Thus equilibrium beliefs are $\Pr_{b_1}(q_i' | r_i = d) = 1$ if the realized value is q_i' and $\Pr_{b_1}(q_i | r_i = d) = 0$ for all other $q_i \neq q_i'$ and likewise for Firm 1. The off-equilibrium beliefs are $\Pr_{b_1}(q_i^1 | r_i = s) = 1$, and $\Pr_{b_1}(q_i | r_i = s) = 0$ for all $q_i \neq q_i^1$ and likewise for Firm 1. This is an equilibrium because the prices are optimal by construction, the buying strategy is optimal by construction, and the beliefs are both rational and consistent.

Now we will consider different cases of strategy profiles with shrouding and we will show that these cannot be equilibria. Observe that Firm 0's equilibrium profit is strictly increasing in Y and Firm 1's equilibrium profit is strictly decreasing in Y.

Case 1 (Partial Shrouding): Without loss of generality, let us consider Firm 0 partially shrouds meaning that there is a non-empty strict subset of quality for which if $q_i \in \tilde{Q}_i \subset \{q_i^1, ..., q_i^m\}$ then $r_i^0(q_i) = s$ but if $q_i \notin \tilde{Q}_i$ then $r_i^0(q_i) = d$.

Subcase A (Partial Shrouding With Exactly One Type): Now consider the case where there is exactly one quality in $q_i' \in \tilde{Q}_i$ and $q_i' \neq q_i^1$. The consumers will believe $E[q_i|r_i^0 = s] = q_i'$. But if the firm has quality q_i^1 it could profitably deviate by shrouding $r_i^0(q_i^1) = s$ which would increase Y. But then $\tilde{Q}_i$ would have more than one quality type which is a contradiction.

Subcase B (Partial Shrouding With At Least Two Types): Under partial shrouding, due to rational beliefs, with the exception of knife-edge cases, the consumer's posterior expected quality will always differ from the prior expected quality: $E_{b_1}[q_i] \neq E_{b_0}[q_i]$. Consequently, attention will shift to this dimension. The consumers' expected quality given shrouding is

$$E[q_i|r_i^0 = s] = \frac{\sum_{k=1}^{l} \Pr(r_i^0 = s|q_i^k) \Pr(q_i^k) q_i^k}{\sum_{k=1}^{l} Pr(r_i^0 = s|q_i^k) \Pr(q_i^k)}.$$

If Firm 0 has $q_i = \max \{\tilde{Q}_i\}$, by disclosing it can increase Y and get a strictly higher profit, so this cannot be part of an equilibrium.

Case 2 (Full Shrouding): Now, without loss of generality, let us consider Firm 0 fully shrouds, meaning that $r_i^0(q_i) = s$ for all q_i. Then the consumers' beliefs will be $E_{b_1}[q_i] = E_{b_0}[q_i]$. We now consider two subcases, first where the other firm does not fully shroud.

Subcase A (Firm 1 Does Not Fully Shroud) If the other firm does not fully shroud then attention will be shifted to the dimension, because except for knife-edge cases, $w_i(B^1; b_1) \neq w_i(B^1; b_0)$. If $q_i > E_{b_0}[q_i]$ then disclosing can only increase Y and hence profit so this cannot be an equilibrium.

Subcase B (Firm 1 Fully Shrouds) Now suppose Firm 1 fully shrouds as well so attention will not be shifted to the dimension. If $q_i > E_{b_0}[q_i]$ and $q_i > E_{b_0}[k_i]$ then disclosing can only increase Y. So this cannot be an equilibrium either.

We have covered all cases where Firm 0 shrouds. Since the firms are symmetric the same logic applies to Firm 1.

References

Akerlof, George A. and William T. Dickens, “The Economic Consequences of Cognitive Dissonance,” *American Economic Review*, 1982, *72* (3), 307–319.

Averill, James R. and Miriam Rosenn, “Vigilant and nonvigilant coping strategies and psychophysiological stress reactions during the anticipation of electric shock,” *Journal of Personality and Social Psychology*, 1972, *23* (1), 128–141.

Bagwell, Kyle, “The Economics of Advertising: Introduction,” in Mark Armstrong and Rob Porter, eds., *Handbook of Industrial Organization*, Vol. 3, North-Holland: Amsterdam, 2007, chapter 28, pp. 1701–1844.

Banerjee, Abhijit V. and Sendhil Mullainathan, “Limited Attention and Income Distribution,” *American Economic Review: Papers and Proceedings*, 2008, *98* (2), 489–493.

Becker, Gary S. and Kevin M. Murphy, “A Simple Theory of Advertising as a Good or Bad,” *The Quarterly Journal of Economics*, 1993, *108* (4), 941–964.

Becker, Marshall H. and Lois A. Maiman, “Sociobehavioral Determinants of Compliance with Health and Medical Care Recommendations,” *Medical Care*, 1975, *13* (1), 10–24.

Bénabou, Roland and Jean Tirole, “Self-Confidence and Personal Motivation,” *The Quarterly Journal of Economics*, Aug 2002, *117* (3), 871–915.

Bennear, Lori S. and Sheila M. Olmstead, “The Impacts of the ’Right to Know’: Information Disclosure and the Violation of Drinking Water Standards,” March 2008. Forthcoming in Journal of Environmental Economics and Management.

Bish, Alison, Amanda Ramirez, Caroline Burgess, and Myra Hunter, “Understanding why women delay in seeking help for breast cancer symptoms,” *Journal of Psychosomatic Research*, 2005, *58*, 321–326.

Bolton, Patrick and Mathias Dewatripont, *Contract Theory*, MIT Press, 2005.

Bracha, Anat and Donald J. Brown, “Affective Decision Making: A Behavioral Theory of Choice,” November 2007. Working Paper.

Brunnermeier, Markus K. and John A. Parker, “Optimal Expectations,” *American Economic Review*, 2005, *94* (4), 1092–1118.

Caplin, Andrew, "Fear As A Policy Instrument," July 2002. Working Paper.

— **and John Leahy**, "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, 2001, *116* (1), 55–79.

— **and —** , "The Supply of Information by a Concerned Expert," *Economic Journal*, 2004, *114*, 487–505.

— **and Kfir Eliaz**, "AIDS Policy and Psychology: A Mechanism-Design Approach," May 2003. Working Paper.

Carr, Deborah and Dmitry Khodyakov, "End-of-Life Health Care Planning Among Young-Old Adults: An Assessment of Psychosocial Influences," *Journal of Gerontology: Social Sciences*, 2007, *62B*, S135–S141.

Carrillo, Juan D. and Thomas Mariotti, "Strategic Ignorance as a Self-Disciplining Device," *The Review of Economic Studies*, 2000, *67* (3), 529–544.

Chesney, Margaret A. and Ashley W. Smith, "Critical Delays in HIV Testing and Care: The Potential Role of Stigma," *American Behavioral Scientist*, 1999, *42* (7), 1162–1174.

DiMatteo, M. Robin, Kelly B. Haskard, and Summer L. Williams, "Health Beliefs, Disease Severity, and Patient Adherence A Meta-Analysis," *Medical Care*, 2007, *45* (6), 521–528.

Eliaz, Kfir and Ran Spiegler, "Consideration Sets and Competitive Marketing," August 2008. Working Paper.

— **and Rani Spiegler**, "Can anticipatory feelings explain anomalous choices of information sources?," *Games and Economic Behavior*, 2006, *56*, 87–104.

Eyster, Erik and Matthew Rabin, "Cursed Equilibrium," *Econometrica*, 2005, *73* (5), 1623–1672.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg, "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *American Economic Review*, 2006, *96* (4), 1043–1068.

Geanakoplos, John, David Pearce, and Ennio Stacchetti, "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1989, *1*, 60–79.

Grossman, Sanford J., "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 1981, *24* (3), 461–483.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh, "Disclosure to a Credulous Audience: The Role of Limited Attention," October 2002. Working Paper.

Hirshleifer, Jack and John G. Riley, *The Analytics of Uncertainty and Information*, Cambridge University Press, 1992.

Hotz, V. Joseph and Mo Xiao, "Strategic Information Disclosure: The Case of Multi-Attribute Products with Heterogeneous Consumers," January 2006. NBER Working Paper 11937.

Hsee, Christopher K., "Attribute Evaluability," in "Choices, Values, and Frames," Cambridge University Press, 2000, chapter 31, pp. 543–563.

Jin, Ginger Zhe and Phillip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, May 2003, *118* (2), 409–451.

Kahneman, Daniel, "Experienced Utility and Objective Happiness," in "Choices, Values, and Frames," Cambridge University Press, 2000, chapter 37, pp. 673–692.

_ , Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur A. Stone, "Would You Be Happier If You Were Richer? A Focusing Illusion," *Science*, 2006, *312* (5782), 1908–1910.

Karlsson, Niklas, George Loewenstein, and Duane Seppi, "The Ostrich Effect: Selective Attention to Information about Investments," September 2008. Working Paper.

Kőszegi, Botond, "Health anxiety and patient behavior," *Journal of Health Economics*, 2003, *22*, 1073–1084.

_ , "Emotional Agency," *Quarterly Journal of Economics*, 2006, *121* (1), 121–155.

_ , "Utility from Anticipation and Personal Equilibrium," *Economic Theory*, August 2009, *Forthcoming.* Working Paper.

_ and Matthew Rabin, "Reference-Dependent Consumption Plans," *American Economic Review*, 2009, *99* (3), 909–936.

Kreps, David M., *Notes on the Theory of Choice*, Westview Press, 1988.

_ and Evan L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 1978, *46* (1), 185–200.

Lindberg, Nangel M. and David Wellisch, "Anxiety and Compliance Among Women at High Risk for Breast Cancer," *Annals of Behavioral Medicine*, 2001, *23* (4), 298–303.

Loewenstein, George, "Anticipation and the Valuation of Delayed Consumption," *The Economic Journal*, 1987, *97* (387), 666–684.

—, **Ted O'Donoghue, and Matthew Rabin**, "Projection Bias in Predicting Future Utility," *The Quarterly Journal of Economics*, November 2003, pp. 1209–1248.

MacPhaila, Catherine and Catherine Campbell, "'I think condoms are good but, aai, I hate those things': condom use among adolescents and young people in a Southern African township," *Social Science & Medicine*, 2001, *52*, 1613–1627.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay, "Revealed Attention," December 2008. Working Paper.

Mayraz, Guy, "Priors and Desires: A Model of Payoff-Dependent Beliefs," June 2009. Working Paper.

Milgrom, Paul and John Roberts, "Relying on the Information of Interested Parties," *RAND Journal of Economics*, 1986, *17* (1), 18–32.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 1981, *12* (2), 380–391.

Miller, Suzanne M. and Charles E. Mangan, "Interacting Effecs of Information and Coping Style in Adapting to Gynecological Stress: Should the Doctor Tell All?," *Journal of Personality and Social Psychology*, 1983, *45* (1), 223–236.

Murry, Jr., John P., Antoine Stam, and John L. Lastovicka, "Evaluating an Anti-Drinking and Driving Advertising Campaign with a Sample Survey and Time Series Intervention Analysis," *Journal of the American Statistical Association*, 1993, *88* (421), 50–56.

Nosarti, C., T. Crayford, J.V. Roberts, E. Elias, K. McKenzie1, and A.S. David, "Delay in presentation of symptomatic referrals to a breast clinic: patient and system factors," *British Journal of Cancer*, 2000, *82* (3), 742–748.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *The American Economic Review*, Mar 1999, *89* (1), 103–124.

Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura, "Strategic Information Revelation," *Review of Economic Studies*, 1990, *57* (1), 25–47.

Osborne, Martin J. and Ariel Rubinstein, *A Course in Game Theory*, MIT Press, 1994.

Schelling, Thomas C., "The Mind as a Consuming Organ," in David E. Bell, Howard Raiffa, and Amos Tversky, eds., *Decision Making*, Cambridge University Press, 1995, chapter 15, pp. 343–357.

Schkade, David A. and Daniel Kahneman, "Does Living in California Make People Happy? A Focusing Illusion in Judgements of Life Satisfaction," *Psychological Science*, 1998, *9* (5), 340–346.

Sherbourne, Cathy Donald, Ron D. Hays, Lynn Ordway, M. Robin DiMatteo, and Richard L. Kravitz, "Antecedents of Adherence to Medical Recommendations: Results from the Medical Outcomes Study," *Journal of Behavioral Medicine*, 1992, *15* (5), 447–468.

Smith, Adam, *The Theory of Moral Sentiments*, Prometheus Books, 1759.

Smith, Ken R. and Robert T. Croyle, "Attitudes toward Genetic Testing for Colon Cancer Risk," *American Journal of Public Health*, 1995, *85* (1), 1435–1438.

Thaler, Richard H. and H. M. Shefrin, "An Economic Theory of Self-Control," *The Journal of Political Economy*, 1981, *89* (2), 392–406.

Tirole, Jean, *The Theory of Industrial Organization*, MIT Press, 1988.

Wood, Wendy, Jeffrey M. Quinn, and Deborah A. Kashy, "Habits in Everyday Life: Thought, Emotion, and Action," *Journal of Personality and Social Psychology*, 2002, *83* (6), 1281–1297.